THE WARING PROBLEM AND ITS GENERALIZATIONS*

BY L. E. DICKSON

1. *Historical Background.* The past year has witnessed greater advances in the field which is the title of this paper than were made during the whole of its previous long history.

Additive number theory had its origin exactly 300 years ago in Fermat's assertion that he was the first to discover the general theorem that every integer is a sum of three triangular numbers, also is a sum of four squares, also is a sum of five pentagonal numbers, and so on to infinity. Here a triangular number $x(x+1)/2$ counts the number of equal spheres arranged in the form of a triangle with x spheres at the base. The proof which he claimed to possess was never published. Cauchy was the first to publish a complete proof in 1813.

In 1659, Fermat stated that his proof of the 4-square theorem was by descent; namely, any integer is a sum of 4 squares provided a certain smaller integer is such. Recently† there was published such a proof, no detail of which was beyond Fermat. The first published proof was that by Lagrange in 1770.

In the same year, Waring conjectured that every positive integer is a sum of 9 integral cubes, positive or zero, also is a sum of 19 fourth powers, and so on. His assertions were certainly empirical, based on short tables.

About 1772, J. A. Euler (son of the celebrated L. Euler) stated that, in order to express every positive integer as a sum of positive nth powers, at least I terms are necessary, where

$$(1) \qquad I = 2^n + q - 2, \qquad q = [(3/2)^n],$$

where, as usual, the notation $[x]$ denotes the largest integer $\leqq x$. He doubtless knew that we must add I powers 1 or 2^n to obtain the sum $2^n q - 1$. For example, we need 4 squares, 9 cubes, 19 fourth powers, 37 fifth powers, and so on.

---

* A paper delivered at the Tercentenary Conference of Arts and Sciences at Harvard University, September, 1936. Presented to the Society, September 2, 1936.

† Dickson, American Journal of Mathematics, vol. 46 (1924), p. 2.

2. *Asymptotic Theory.* The pioneer work was that by Hardy and Littlewood in 1919–1927. Let $G(n)$ denote the least integer s such that every sufficiently large integer is a sum of s integral nth powers, each $\geqq 0$. Their initial result was

$$G(n) \leqq (n-2)2^{n-1} + 5,$$

whence it suffices to use 9 cubes, 21 fourth powers, 53 fifth powers, and so on. In 1925, they found for $n \geqq 4$ the lower result

$$G(n) \leqq (n-2)2^{n-2} + n + 5 + \zeta_n, \tag{2}$$

where ζ_n is a relatively small function $< n^2 \log_e 2$. Their earlier 53 fifth powers are now replaced by 41.

In 1934, Vinogradow* gave $G(n) < 32(n \log_e n)^2$. He† soon obtained the astonishing result that, if $n \geqq 4$,

$$G(n) \leqq 4n + 3k, \qquad k = [2n \log_e n + n \log_e 6]. \tag{3}$$

This exceeds the former result given by (2) if $n = 4$, 5, or 6, but is a decided improvement if $n \geqq 7$. For example, if $n = 17$, the formula (3) gives 448, in contrast with the value 491711 given by (2).

Later I supplied details‡ for Vinogradow's proof, evaluated all the constants, and proved that if

$$\log_{10} \log_{10} N > 0.6621n^2, \qquad (n \geqq 10), \tag{4}$$

every integer $\geqq N$ is a sum of nth powers whose number is given by (3).

3. *Ideal Waring Theorem.* The ideal Waring theorem states that every positive integer is a sum of I integral nth powers $\geqq 0$, for I in (1), or technically, $g(n) = I$.

Early in 1935, I proved by use of (4) the ideal Waring theorem when $n = 11$, 12, 13, 14, 15, 17, while Zuckerman proved it independently when $n = 14$-20. When $n \geqq 21$, I and hence $g(n)$ exceeds 2 million, so that there seems to be little satisfaction in knowing the exact value of the large number $g(n)$, which is already known to be finite.

* Comptes Rendus de l'Académie des Sciences, URSS, vol. 2 (1934), pp. 337–341.

† Annals of Mathematics, vol. 36 (1935), pp. 395–405.

‡ Annals of Mathematics, vol. 37 (1936), pp. 293–316. With extension to sums of nth powers multiplied by any given positive integers.

4. *Solution of Waring's Problem.* An essential step was a material lowering of the constant given by (4). The largeness of the constant given by (4) was due mainly to the approximation to the number of the divisors of an integer. This divisor function is entirely avoided in a new treatment* of the final part of the asymptotic theory, which shows that, if

$$\log_{10} N = 2n^7, \qquad (n \geqq 9), \tag{5}$$

every integer $\geqq N$ is a sum of $M=4n-2+6k$ integral nth powers. Here $M<I$. It remains to prove that every positive integer $<N$ is a sum of I nth powers. To indicate the general nature of this proof, note that if every integer J between E and $E+2^n$ is known to be a sum of m nth powers, then every integer between $E+2^n$ and $E+2\cdot 2^n$ can be expressed in the form $J+2^n$ and hence is a sum of $(m+1)$ nth powers. We readily compute m when $E=2^nq$, so that our initial interval contains 3^n. Hence the preceding fact shows that all integers between 2^nq and 2^nq+3^n are sums of $m+q$ nth powers. Then all integers on to $2^nq+2\cdot 3^n$ are sums of $m+q+1$ powers, and so on. After making a certain number of such single ascents, we reach a point for which we may make a general number of ascents at one step by use of a general formula† involving logarithms. In this way I obtained the following result.

THEOREM 1. *Let* $3^n=2^nq+r$, $0<r<2^n$. *If* $n>6$ *and*

$$2^n \geqq q + r + 3, \tag{6}$$

then $g(n)=I$ (*that is, the ideal Waring theorem holds*).

This inequality holds if $4\leqq n\leqq 400$. If it should fail for any $n>400$, then the decimal for $r/2^n$ would begin with at least 50 figures 9. Below 400, it begins with two 9's only when $n=157$ and $n=163$, and for both the third figure is less than 9.

Although (6) will rarely, if ever, fail, I have proved also the following fact.

THEOREM 2. *If* (6) *fails, write* $f=[(4/3)^n]$. *Then* $g(n)=I+f$ *or* $I+f-1$, *according as* $2^n=$ *or* $<fq+f+q$.

* Dickson, American Journal of Mathematics, vol. 58 (1936), pp. 521–535; announced March 13 in this Bulletin, vol. 42 (1936), p. 341.

† Dickson, this Bulletin, vol. 39 (1933), p. 711.

Theorem 1 for $n \geqq 8$, but not Theorem 2, was obtained independently by S. S. Pillai.* Now that Waring's problem has been solved when $n \geqq 6$, it is natural to undertake the following two direct generalizations.

5. *First Generalization.* Define $g(n, m)$ so that every integer $\geqq m$ is a sum of g nth powers, while not every integer $\geqq m$ is a sum of $g-1$ powers. Thus $g(n, 0) = g(n, 1) = g(n)$.

While $g(12) = 4223$, I proved† that $g(12, 2 \cdot 3^{12}) = 2405$, $g(12, 3 \cdot 3^{12}) = 1560$, $g(12, 8^{12}+7^{12}+2 \cdot 5^{12}) \leqq 440$. Also,‡ $g(14, 4^{14}) = 5184 < I/3$, and so on. I obtain§ many such results for $n = 9$ and $n = 11$.

6. *Second Generalization.* I shall now announce‖ remarkable new results.

THEOREM 3. *Let $d = 1$ or 2, according as q is odd or even. If $9 \leqq n \leqq 400$, every positive integer is a sum of $4n+2-d$ integral nth powers $\geqq 0$ and the doubles of $P = \frac{1}{2}(2^n+q-4n+d)-2$ such powers. Here $4n+2-d+2P = I$. Expressed otherwise, in the ideal Waring theorem we may take $2P$ of the powers equal in pairs.*

We have therefore reduced the number of variables or summands from I to $I-P$. Except for small n's, $I-P$ just exceeds $I/2$.

Similar theorems reduce the number of summands from I to approximately I/a, for any integer a that is not too large.

THEOREM 4. *If $a \leqq 4n$, $16 \leqq n \leqq 400$, every positive integer is a sum of $m = 4n+2an$ integral nth powers $\geqq 0$ and the a-tuples of $\{(I-m)/a\}-2$ powers, where $\{x\}$ denotes the least integer $\geqq x$.*

THEOREM 5. *If $a = 3$ or 4, $11 \leqq n \leqq 400$, Theorem 4 holds with $m = 4n$.*

Evident corollaries are obtained from Theorems 4 and 5 by partitioning each a into arbitrary positive integral parts, which may differ for the various a's. For example, $5x^n = 3x^n + 2x^n$.

---

* Journal of the Indian Mathematical Society, vol. 2 (1936), pp. 16–44.

† Duke Mathematical Journal, vol. 2 (1936), pp. 192–204.

‡ Monatshefte für Mathematik und Physik, vol. 43 (1936), pp. 391–400.

§ This Bulletin, vol. 42 (1936), pp. 525–529.

‖ To appear in volume 2 of Acta Arithmetica.

In the proof of these theorems for the small values of n, it was necessary to resort to facts obtained by tables. It is an interesting and economical fact that we need not construct tables of sums of values of these new functions involving a, but that we may derive the necessary information from published tables of sums of values of x^n. In fact, I prove the following theorem.

THEOREM 6. *For $g^n \leqq E < 2g^n$, let all integers in the interval $K=(E, E+P)$ be sums of h integral nth powers $\geqq 0$. Then all integers in the same interval K are sums of $(g-1)(a-1)+1$ integral nth powers and the a-tuples of $[h/a]$ powers.*

7. *Polynomial Summands (Asymptotic Results).* Without precision or any details, Waring implied that there is a Waring problem with polynomial summands $f(x)$. Kamke* employed a polynomial $f(x)$ of degree $n \geqq 2$ having rational coefficients, that of x^n being positive, such that $f(x)$ is an integer $\geqq 0$ for every integer $x \geqq 0$, and proved that every positive integer is a sum of a limited number of 1's and a limited number of values of $f(x)$ for integers $x \geqq 0$. The proof rests on the simultaneous solvability of the equation

$$l_j = x_1^j + \cdots + x_N^j, \qquad (j = 1, \cdots, n), \tag{7}$$

in integers $x \geqq 0$, when the l_j are given integers lying in certain intervals, while N is an existing function of n. It would be interesting to know the minimum value of N.

Equations (7) for $n=2$, also when coefficients are inserted, have been recently† discussed at length.

For

$$s \geqq (n-2)2^{n-1} + 5,$$

Landau‡ found an asymptotic expression for the number of ways to express a given integer as a sum of s values of $f(x)$.

Further progress was blocked by the lack of a solution of the following problem: To find a reasonably small value of $m=m(f, p)$ such that every integer is congruent modulo p (a

* Mathematische Annalen, vol. 83 (1921), pp. 85–112. Journal für Mathematik, vol. 152 (1923), pp. 30–32.

† Dickson, American Journal of Mathematics, vol. 56 (1934), pp. 513–528; Quarterly Journal of Mathematics, vol. 5 (1934), pp. 283–290.

‡ Mathematische Zeitschrift, vol. 31 (1930), pp. 319–338.

prime) to a sum of m values of the polynomial $f(x)$. For any f of degree $n<28$, I proved* that we may take m smaller than the value given by the equation (2) of Hardy and Littlewood.

Under specified mild restrictions on the coefficients of $f(x)$, R. D. James† proved that $G(f)\leqq 9$ if $n=3$, while Miss M. G. Humphreys‡ proved that

$$G(f) \leqq (n-2)2^{n-1}+5$$

if $n=4$, 5, 6, or 7, and under strong conditions on $f(x)$ also for $8\leqq n\leqq 28$.

K. S. Ghent§ extended Landau's asymptotic result to the expression

$$F = a_1 f(x_1) + \cdots + a_s f(x_s), \tag{8}$$

where the a_i are given positive integers. Under mild conditions on the a_i and f, he proved that $G(F)\leqq 9$ for $n=3$ and that $G(F)\leqq 21$ for $n=4$.

L. Hua writes me that he has proved that $G(f)$ is less than or equal to the value given by (2) if f is an odd polynomial such that, for no integers l and d $(d>1)$, $f(x)\equiv l \pmod d$ for every x.

The preceding papers found for $G(f)$ a limit which is the same as the initial or final limit for $G(x^n)$ obtained by Hardy and Littlewood. It is hoped that future investigations will reduce $G(f)$ to a limit at least close to that obtained by Vinogradow for $G(x^n)$. We should then be able to evaluate $g(f)$.

8. *Value of $g(f)$ for Polynomials f.* For the polynomial

$$f(x) = \frac{1}{6}\epsilon(x^3-x)+\sigma x, \qquad \sigma \not\equiv 2a \pmod 3 \text{ if } \epsilon = 3a, \tag{9}$$

we can actually determine‖ positive integers C and v such that every integer $\geqq N = C\cdot 3^{3v}$ is a sum of nine values of (9). Here N is relatively small compared to the constant arising from the

* American Journal of Mathematics, vol. 57 (1935), pp. 463–474. Proof by cyclotomy.

† American Journal of Mathematics, vol. 56 (1934), pp. 303–315.

‡ Duke Mathematical Journal, vol. 1 (1935), pp. 361–375.

§ Chicago Dissertation, 1935, 49 pp. Condensed in Duke Mathematical Journal.

‖ Dickson, Transactions of this Society, vol. 36 (1934), pp. 1–12, 739–741.

analytic theory, whence the algebraic theory is the better for finding $g(f)$. If $\sigma=1$, $1\leqq\epsilon\leqq 6$, then* $g(f)\leqq 9$. A. Sugar readily computed $g(f)$ for any ϵ when $\sigma=1$.

Elsewhere I obtained† systematically 116 cubic polynomials F for which $g(F)\leqq 9$. For many of them, $g(F)=4$ or 5 empirically.

The analytic theory of prime numbers has been applied‡ to show that all large numbers are sums of 8 values of (9) for certain sets ϵ, σ.

Given a Waring theorem for a polynomial $q(x)$ of degree n, we may deduce§ instantaneously a Waring theorem for $f(x)-hx$, where f is an even function of degree $2n$ and h is any integer. Hence we deduce an upper limit to $g(F)$, for many polynomials F of degrees 4, 6, 8, 10.

9. *A Paradox.* Granted that an integer m is a sum of s integral values of the function x^n, for example, we see that $m+sc$ is evidently a sum of s values of x^n+c, where c is a constant positive integer, and conversely. It would therefore seem that a Waring problem for x^n is the same as for x^n+c, except for small integers. While this is true in the above unimportant sense, it is false in the really important case.

For summands x^n, it is immaterial whether we say in modern parlance that m is a sum of exactly s values (one of which is zero), or whether we return to the periods of Fermat, Waring, and others, and say that m is a sum of s or fewer (positive) values. But the function x^n+c does not take the value zero for an integral value $\geqq 0$ of x and the corresponding two problems are essentially different.

I have proved‖ quite recently that, if $n\geqq 9$,

$$g(x^n+1)=2^{n-1}+q-1,$$

* By $g(f)=m$ we mean that all positive integers are sums of m values of f, but not all are sums of $m-1$ values of f.

† Acta Arithmetica, vol. 1 (1936), pp. 184–196.

‡ James, Mathematische Annalen, vol. 109 (1933), pp. 196–199. L. Hua, ibid., vol. 111 (1935), pp. 622–628.

§ Dickson, Transactions of this Society, vol. 36 (1934), pp. 371–378; vol. 39 (1936), pp. 205–208.

‖ Here $g(f)=m$ means that all positive integers are sums of m or fewer values of f, but not all are sums of $m-1$ or fewer values.

which is materially smaller than I, where

$$q = [(3^n + 1)/(2^n + 1)].$$

Also that $g(x^3+1) \leqq 8$. But proofs are not yet possible for the best results for small n's. I find that all positive integers $< N$ are sums of k or fewer values of x^n+1, where the following table holds:*

n	3	4	5	6	7
$\log_{10} N$	9	4.6115	27.616	38.8893	4405.26
k	7	14	22	41	85
I	9	19	37	73	143

For $f(x) = x^n+2$, let $Q = [f(3)/f(2)]$ and let $\{x\}$ denote the least integer $\geqq x$. For $n \geqq 9$, every integer >1 is a sum of $Q-1+\{(2^n+3)/3\}$ or fewer values of x^n+2.

For $n \geqq 11$, every integer >2 (except 5) is a sum of $Q+2^{n-2}+1$ or fewer values of x^n+3.

n	9	11	12	14
$g(x^n)$	548	2132	4223	16673
$g(x^n+1)$	293	1109	2176	8482
$g(x^n+2, 1)$	209	769	1495	5753
$g(x^n+3, 5)$	167(?)	599	1154	4338

All integers between 1 and one billion are sums of 7 or fewer values of x^3+2. Further reductions of g arise when $c \geqq 4$. But now various small integers are not sums of values of x^n+c.

As in §6, we may take all but s, $(s \geqq 4n)$, of the summands equal in sets of a. For example, let $c=1$, $a=3$. Write $m = 2n/3$ if n is divisible by 3, but $m = 1+2[n/3]$ if n is prime to 3. Then every positive integer is a sum of $3+6m$ or fewer values of x^n+1 and the triples of

$$(2^{n-1} \pm 1)/3 + [(q-1)/3] - m$$

or fewer values, where the sign is $+$ or $-$ according as n is even or odd.

* For $n=4$, only 487–488 require 14. For $n=5$, only 229–230 and 961–962 require 22. For $n=6$, only 713–714 require 41. We made only a short new table to reach a point beyond which known tables for x^n are applicable.

10. *Homogeneous Summands.* A highly important feature of Vinogradow's method is the use of a parameter y^n. If we replace his summand x^n by $F(x, z)$ we must have $y^nF(x, z) = F(X, Z)$, whence F must be homogeneous and of degree n. In this case, H. Chatland has overcome the difficulties arising by the replacement of x^n by F except the following ones. We seek forms F with many distinct values. To be precise, let F take M distinct values when $0 \leqq x \leqq Q$, $0 \leqq z \leqq Q$; we desire that $M/(Q+1)^2$ shall be 0.99 or some number just less than unity. I would welcome any information.

I have treated this problem for forms divisible by the square or higher power of a linear form. Let $f(x, z)$ be a form of degree t whose coefficients are integers $\geqq 0$, that of z^t being greater than zero. We obtain only distinct values of $F = x^m f$ if we employ integers $x > 0$, $y \geqq 0$ such that f is not divisible by any mth power > 1. It suffices to treat $f \equiv 0 \pmod{p^m}$, p a prime, and hence to discuss $z \equiv cx \pmod{N}$, where N is a power of p. The latter problem would be the trivial classic one if x, z took the values $0, 1, \cdots, N-1$. On the contrary, they shall here take values $\leqq Q$. The problem is solved by use of the following theorem.

Let c and N be given integers, $0 < c < N$. Let X denote the least residue $\geqq 0$ of cx modulo N. Let r be the greatest integer $< N/c$. Then, for $1 \leqq x \leqq Q$, there are at most $[Q/r]$ values $X < c$.

Evidently we shall obtain a lower* value for $g(F)$ than for the case of a function of one variable.

For

$$F(x, z) = x^m(z^n + cx^n),$$

we have

$$F(1, z) = z^n + c.$$

Hence $g(F)$ is smaller than $g(z^n+c)$, which was found in §9. In particular, all positive integers $\leqq 12466$ are sums of 13 or fewer values of $x^4(x^4+z^4)$.

James† found $g(x^3+cxy^2)$ for every c. The case $y=0$ shows that $g \leqq 9$.

* Dickson, American Journal of Mathematics, vol. 48 (1936), pp. 241–248.

† Journal of the London Mathematical Society, vol. 10 (1935), pp. 84–88.

11. *Latest Asymptotic Papers.* H. Heilbronn* gave a much shorter proof of the result by Vinogradow; but it leads to a far greater constant.

Vinogradow† modified his method by using Weil's approximation and obtained $G(n) \leqq 2[n(n-2)\log_e 2+2n]$, which is 40, 56, 214 when $n=5, 6, 12$.

Proofs that $G(4) \leqq 17$ were given independently by T. Estermann‡ and by H. Davenport and Heilbronn.§ The method used by the latter has been extended to any exponent n by Alaoglu, who found $G(5) \leqq 30$, $G(6) \leqq 51$.

For the writer's report on Waring's problem see an article in this Bulletin (vol. 39 (1933), pp. 701–727). On the converse problem, see another article in this Bulletin (vol. 40 (1934), pp. 711–714).

THE UNIVERSITY OF CHICAGO

---

* Acta Arithmetica, vol. 1 (1936), pp. 212–221.

† Mathematical Institute, Stekloff, vol. 9 (1935), pp. 5–15. (Russian.) Announced in Comptes Rendus, Paris, vol. 199 (1934), pp. 174–175.

‡ Proceedings of the London Mathematical Society, vol. 41 (1936), pp. 126–142.

§ Proceedings of the London Mathematical Society, vol. 41 (1936), pp. 143–150, 449–453.

NOTE ON DIVISIBILITY SEQUENCES

BY MORGAN WARD

1. *Introduction.* We call a sequence of rational integers

$$(u): \qquad u_1, u_2, u_3, \cdots, u_n, \cdots$$

a *divisibility sequence* if u_r divides u_s whenever r divides s. The divisibility sequences most frequently studied are the *linear* sequences which satisfy linear difference equations with constant, integral coefficients.* In particular, the divisibility sequence associated with a difference equation of order two is essentially one of the important functions of Lucas.† I propose here to deduce two striking properties of divisibility sequences which do not depend on the fact that the sequence is linear.

2. *Preliminary Definitions.* An integer m will be said to be a *divisor* of (u) if it divides some term of (u), and a *prime divisor* if it is a prime. The suffix of the first term of (u) divisible by m is called the *rank of apparition* of m. If p is a prime divisor of (u), the rank of apparition of p^a, if it exists, will be denoted by ρ_a.

If we assume that no term of (u) is zero, we can build up from (u) a set of numbers $[n, r]$, the *binomial coefficients belonging to* (u),‡ defined by

$$[n, r] = 1, \qquad (r = 0; n = 0, 1, 2, \cdots),$$

$$[n, r] = u_n \cdot u_{n-1} \cdot \cdots \cdot u_{n-r+1} / u_1 \cdot u_2 \cdot \cdots \cdot u_r, \qquad (r = 1, \cdots, n; n = 1, 2, \cdots).$$

They will not in general be rational integers.

If a and b are any rational integers, we shall write as usual $a \mid b$ for a divides b and (a, b) for the greatest common divisor of a

---

* See Marshall Hall, *Divisibility sequences of the third order*, American Journal of Mathematics, vol. 58 (1936), pp. 577–584, for an account of these sequences and references to the work of Pierce, Poulet, and Lehmer.

† u_n equals the function $(\alpha^n - \beta^n)/(\alpha - \beta)$ up to a constant factor.

‡ For a systematic account of the remarkable properties of these numbers formed from any sequence (u) with no non-vanishing terms see Morgan Ward, *A calculus of sequences*, American Journal of Mathematics, vol. 58 (1936), pp. 255–266.

and b. If a^r is the highest power of a which divides b, we shall write $a^r \| b$.

Finally, since u_1 must divide every term of (u), we may assume that $u_1 = 1$.

3. *Statement of Results.* A divisibility sequence will be said to have property A provided that

A. If $c = (a, b)$, then $u_c = (u_a, u_b)$, for every pair of terms u_a, u_b of (u).

It will be said to have property B provided that

B. For every prime divisor p and every positive integer a, $u_r \equiv 0 \pmod{p^a}$ when and only when $r \equiv 0 \pmod{\rho_a}$, where ρ_a is the rank of apparition of p^a in (u).

The results of this note may now be stated as follows.

THEOREM 1. *Property* A *and property* B *are equivalent to one another.*

THEOREM 2. *The binomial coefficients belonging to every divisibility sequence having property* A *or property* B *are all rational integers.*

Theorem 2 was proved for the Lucas function by Lucas himself,* and for a more general type of linear divisibility sequence by T. A. Pierce.†

4. *Proof of First Theorem.* Assume that the divisibility sequence (u) has property A, and let ρ_a be the rank of apparition of p^a, where p is any prime divisor of (u). Suppose that $u_r \equiv 0 \pmod{p^a}$. Then if $c = (r, \rho_a)$, $(u_r, u_{\rho_a}) = u_c$ by property A. Therefore since $u_r \equiv u_{\rho_a} \equiv 0 \pmod{p^a}$, $u_c \equiv 0 \pmod{p^a}$. Therefore $c \geqq \rho_a$. But c divides ρ_a. Therefore $c = \rho_a$ so that ρ_a divides r. Since (u) is a divisibility sequence, if ρ_a divides r, $u_r \equiv 0 \pmod{p^a}$. Therefore the sequence has property B.

Conversely, assume that (u) has property B. Let u_a and u_b be any two terms of (u), and let p be any common prime divisor of u_a and u_b. Suppose that $p^m \| u_a$ and $p^n \| u_b$. Then if l is the smallest of the integers m and n, *it suffices to show that* $p^l \mid u_c$, *where* $c = (a, b)$. For since $c \mid a$ and $c \mid b$, $u_c \mid u_a$ and $u_c \mid u_b$, so that

* Lucas, Nouvelle Correspondance Mathématique, vol. 4 (1878), pp. 1–8. Dickson's *History*, vol. 1, p. 349.

† Annals of Mathematics, (2), vol. 18 (1916–17), p. 56.

$u_c \mid (u_a, u_b)$. But $p^l \| (u_a, u_b)$. Therefore if $p^l \mid u_c$ for every common prime divisor p of u_a and u_b, we have $(u_a, u_b) \mid u_c$, so that $(u_a, u_b) = u_c$, and property A follows.

Now let ρ_m, ρ_n be the ranks of apparition of p^m and p^n, respectively. Without loss of generality we may assume that $m \geqq n$, so that $l = n$. Since property B holds, $\rho_m \mid a$, $\rho_n \mid b$ and $\rho_n \mid \rho_m$. Hence $\rho_n \mid a$ and $\rho_n \mid b$, so that $\rho_n \mid c = (a, b)$. But then $u_{\rho_n} \mid u_c$, so that $p^l = p^n \mid u_c$.

5. *Proof of Second Theorem.* It suffices to show that $[n, r]$ is an integer modulo p for every prime divisor p of (u) when (u) has property B. If we let $[0]! = 1$, then $[s]! = u_1 u_2 \cdots u_s$, $(s \geqq 1)$, $[n, r] = [n]!/[n-r]![r]!$.

Now the highest power of p dividing $[n]!$ is clearly $\sum_{s=1}^{\infty} [n/\rho_s]$, where as usual $[a/b]$ denotes the greatest integer in a/b. (If p^s does not divide (u), then neither does p^t, $(t \geqq s)$, and we break off the sum after $s-1$ terms. Since $\rho_s \to \infty$ with s if every power of p divides the sequence, the sum is finite in every case.)

It therefore suffices to show that

$$\sum_{s=1}^{\infty} \left[\frac{n}{\rho_s}\right] \geqq \sum_{s=1}^{\infty} \left[\frac{n-r}{\rho_s}\right] + \sum_{s=1}^{\infty} \left[\frac{r}{\rho_s}\right],$$

and this follows as in the ordinary case when $u_n = n$ from the elementary inequality

$$\left[\frac{n+m}{\rho}\right] \geqq \left[\frac{n}{\rho}\right] + \left[\frac{m}{\rho}\right].$$

CALIFORNIA INSTITUTE OF TECHNOLOGY

A SIMPLIFIED SET OF POSTULATES FOR A GROUP*

BY R. M. FOSTER

Let there be given a class of elements $(a, b, c, \cdots)$ denoted by G, and an operation by which any two ordered elements determine a unique result which may or may not be an element of the class G. For convenience, this operation may be called multiplication, and the result determined by a and b called the product and written ab. It is the purpose of the present note to show that this system then forms a group if it satisfies the following set of three postulates:

I. *If a, b, c, ab, bc, $(ab)c$, and $a(bc)$ are all elements of G, then $(ab)c=a(bc)$.*

II. *If a and b are elements of G, there exists an element x of G such that $ax=b$.*

III′. *There exists an element g of G with the property that, if b is an element of G, there exists an element y of G such that $yg=b$.*

It has been shown by Garver† that a finite group may be defined by these three postulates together with a fourth:

IV. *The number of elements of G is finite.*

To define a general group, however, Garver in an earlier paper‡ had required a set of three postulates consisting of I, II, and a stronger form of III′, namely,

III. *If a and b are elements of G, there exists an element y of G such that $ya=b$.*

Whereas Garver showed that in defining a finite group by the four postulates I, II, III, and IV, a part of III is redundant and that a considerable simplification can be effected upon re-

* Presented to the Society, October 31, 1936.

† Raymond Garver, *Postulates for special types of groups*, this Bulletin, vol. 42 (1936), pp. 125–129.

‡ Raymond Garver, *Note concerning group postulates*, this Bulletin, vol. 40 (1934), pp. 698–701.

placing III by III′, it is shown in the present note that the same simplification can be effected in the set of postulates for a general group—that Garver's simplification is not essentially limited to the special case of finite groups. In other words, postulates I, II, and III′ suffice to define a group; and thus, without any revision, I, II, III′, and IV define a finite group.

The method of proof consists in establishing the existence of an identity element, and then of an inverse for each element; thereupon it is shown that the system has the closure property (if a and b are elements of G, then ab is an element of G). It is convenient to arrange the first part of this proof as the deduction of the three following lemmas.

LEMMA 1. *There exists an element e of G such that $ea=a$ for every element a of G.*

By III′ there exist elements g and e of G such that $eg=g$. If a is an element of G, then, by II, there exist elements m and n of G such that $em=a$ and $gn=m$. Thus $m=gn=(eg)n$, and $a=em=e(gn)$. Hence by I, $m=a$; and upon substituting in $em=a$, we have $ea=a$ for every element a.

LEMMA 2. *If e is the element of G having the property of Lemma 1, then $ae=a$ for every element a of G.*

By III′ there exists an element p of G such that $pg=e$. If a is an element of G, then, by II, there exist elements q, r, and s of G such that $aq=a$, $pr=q$, and $gs=r$. By Lemma 1, $s=es$. Thus $s=(pg)s$, and $q=pr=p(gs)$. Hence by I, $s=q$, and upon substituting in $gs=r$, we have $gq=r$.

By III′ there exists an element t of G such that $tg=a$. By II, there exist elements u and v of G such that $tu=e$ and $uv=g$. By Lemma 1, $v=ev$. Thus $v=(tu)v$, and $a=tg=t(uv)$. Hence, by I, $v=a$; and upon substituting in $uv=g$, we have $ua=g$.

We now have $r=gq=(ua)q$, and $g=ua=u(aq)$. Hence by I, $r=g$, and thus $gq=g$.

By Lemma 1, $q=eq$. Thus $q=(pg)q$, and $e=pg=p(gq)$. Hence by I, $q=e$, and upon substituting in $aq=a$, we have $ae=a$ for every element a of G.

LEMMA 3. *If e is the element of G having the properties of Lemmas 1 and 2, and if a is an element of G, there exists an element a^{-1} of G such that $aa^{-1}=a^{-1}a=e$.*

By II there exist elements a^{-1} and w of G such that $aa^{-1}=e$ and $a^{-1}w=e$. By Lemma 1, $w=ew$; and by Lemma 2, $a=ae$. Thus $w=(aa^{-1})w$, and $a=a(a^{-1}w)$. Hence by I, $w=a$, and, upon substituting in $a^{-1}w=e$, we have $a^{-1}a=e$.

We are now in a position to establish the following closure property.

V. *If a and b are elements of G, there exists an element z of G such that $ab=z$.*

By II and the lemmas, there exist elements e, a^{-1}, z, and k such that $aa^{-1}=e$, $a^{-1}z=b$, and $ak=z$. Then by Lemma 1, $k=ek$; and by Lemma 3, $e=a^{-1}a$. Thus $k=(a^{-1}a)k$, and $b=a^{-1}z=a^{-1}(ak)$. Hence by I, $k=b$, and upon substituting in $ak=z$, we have $ab=z$, thus establishing the closure property.

The proof of III then follows by a familiar derivation from closure and the lemmas established above. Since I, II, and III define a group, as shown by Garver in his 1934 paper, it thus follows that I, II, and III′ define a group.

It is ordinarily assumed that multiplication is unique, and this assumption is not made one of the explicit postulates. For some purposes it might be better to incorporate uniqueness as one of the postulates of the set; in this case certain modifications would be desirable in the statement of the associative law. This procedure was followed by Baer and Levi* in a comprehensive study of group postulates. Their paper also gives an answer to the question raised by Garver at the end of his 1936 paper, showing by means of an example that it is not possible to define a commutative group by means of postulates I and II and the weak form of the commutative law (if a, b, ab, and ba are all elements of G, then $ab=ba$).

BELL TELEPHONE LABORATORIES

* Reinhold Baer and Friedrich Levi, *Vollständige irreduzibele Systeme von Gruppenaxiomen*, Sitzungsberichte der Heidelberger Akademie der Wissenschaften, Mathematisch-naturwissenschaftliche Klasse, Beitrag zur Algebra, Nr. 18 (1932), pp. 3–12.

ON THE GENERATION OF THE FUNCTIONS Cpq AND Np OF LUKASIEWICZ AND TARSKI BY MEANS OF A SINGLE BINARY OPERATION

BY J. C. C. McKINSEY*

Indicating the n "truth-values" of a Lukasiewicz-Tarski logic† by the n numbers 1, 2, $\cdots$, n, we define the functions Cpq and Np as follows:

$Cpq = 1$, when $p \geqq q$,
$Cpq = q - p + 1$, when $p < q$,
$Np = n - p + 1$.

Thus, for example, for $n = 3$ we have

C	1	2	3
1	1	2	3
2	1	1	2
3	1	1	1

p	Np
1	3
2	2
3	1

I shall denote a Lukasiewicz-Tarski logic of n truth-values by L_n.

In this paper I define,‡ in terms of Cpq and Np, a function E_ipq such that, in each L_n, Cpq and Np are in turn definable in terms of $E_{n-2}pq$. The function E_ipq is defined by means of the following series of definitions.

DEFINITION 1. $A_0p = p$, $A_{i+1}p = CNpA_ip$.
DEFINITION 2. $B_0p = Np$, $B_{i+1}p = CpB_ip$.
DEFINITION 3. $D_ip = CA_ipNCpNB_ip$.
DEFINITION 4. $E_ipq = CpD_iq$.

---

* Blumenthal Research Fellow.

† For a general discussion of this logic, see Lewis and Langford, *Symbolic Logic*, pp. 199–234.

‡ D. L. Webb has recently found (*The generation of any n-valued logic by one binary operation*, Proceedings of the National Academy of Sciences, vol. 21 (1935), pp. 252–254) a binary operation by means of which it is possible to generate any operation of any n-valued logic. His operation, however, cannot be defined in terms of Cpq and Np except when $n = 2$. This can be seen from the fact that the operations Cpq and Np are class-closing on the elements 1, n; whereas the operation found by Webb has not this property.

In terms of E_ipq I define certain other functions as follows:

DEFINITION 5. $F_ip = E_iE_ippE_iE_ippE_ipp$.
DEFINITION 6. $M_ip = E_ipF_ip$.
DEFINITION 7. $I_ipq = E_ipE_iF_iqq$.

I shall now show that, in L_n, $M_{n-2} = Np$, and $I_{n-2}pq = Cpq$; hence that, in L_n, Cpq and Np are definable in terms of the single binary operation $E_{n-2}pq$.

THEOREM 1. *For every n in L_n we have*

$$A_{n-2}n = n, \quad \text{and} \quad A_{n-2}p = 1 \text{ for } p \neq n.$$

PROOF. I prove the first part of the theorem by mathematical induction on i. By Definition 1, $A_0n = n$. Suppose that $A_kn = n$; then $A_{k+1}n = CNnA_kn = CNnn = Cln = n$. Hence for every i we have $A_in = n$; so, in particular, $A_{n-2}n = n$.

I prove the second part of the theorem by *reductio ad absurdum*. Suppose, if possible, that the second part of the theorem is false, so that there exists a $p_0 < n$ for which $A_{n-2}p_0 > 1$. I first show that, on this supposition, $A_ip_0 > 1$ for every $i \leqq n-2$; for if we had $A_ip_0 = 1$ we should have $A_{i+1}p_0 = CNp_0A_ip_0 = CNp_01 = 1$, so we should have $A_{n-2}p_0 = 1$, contrary to hypothesis. It can be shown that $A_1p_0 \leqq n-2$; for from $p_0 < n$ follows $p_0 \leqq n-1$, whence $2p_0 \leqq 2n-2$, whence $2p_0 - n \leqq n-2$; and, since $A_1p_0 \neq 1$, we have $A_1p_0 = CNp_0p_0 = p_0 - (Np_0) + 1 = p_0 - (n - p_0 + 1) - 1 = 2p_0 - n$. It can also be shown that for each k, $(n-2 > k > 1)$, we have $A_{k+1}p_0 < A_kp_0$; for from $p_0 < n$ follows $n - p_0 + 1 > 1$, so $Np_0 > 1$; whence $A_kp_0 - Np_0 + 1 < A_kp_0$, and since $A_{k+1}p_0 \neq 1$, we have $A_{k+1}p_0 = A_kp_0 - Np_0 + 1$. Thus we have

$$A_{n-2}p_0 < A_{n-3}p_0 < \cdots < A_2p_0 < A_1p_0 \leqq n - 2.$$

Hence

$$A_{n-2}p_0 \leqq A_1p_0 - (n-3) \leqq (n-2) - (n-3),$$

and $A_{n-2}p \leqq 1$. But this is contrary to hypothesis. Hence the second part of the theorem is true.

The proof of the following theorem is similar.

THEOREM 2. *For every n in L_n we have*

$$B_{n-2}1 = n, \quad \text{and} \quad B_{n-2}p = 1 \text{ for } p \neq 1.$$

THEOREM 3. *For every n in L_n we have*

$$D_{n-2}1 = n, \qquad D_{n-2}n = 1, \qquad D_{n-2}p = p \text{ for } p \neq 1, n.$$

PROOF. By Theorems 1 and 2, and the definitions of Cpq and Np, we have $D_{n-2}1 = CA_{n-2}1NC1NB_{n-2}1 = C1NC1Nn = C1NC11 = C1n = n$, $D_{n-2}n = CA_{n-2}nNCnB_{n-2}n = CnNCn1 = CnNn = Cn1 = 1$. Suppose now that $p \neq 1, n$. Then $D_{n-2}p = CA_{n-2}pNCpNB_{n-2}p = C1NCpN1 = C1NCpn = C1N(n-p+1) = C1[n-(n-p+1)+1] = C1p = p$.

THEOREM 4. *For every $p \neq 1$ in L_n, $E_{n-2}pp = 1$; and $E_{n-2}11 = n$.*

PROOF. If $p \neq 1, n$ then, by Theorem 3, $E_{n-2}pp = CpD_{n-2}p = Cpp = 1$. If $p = n$, then $E_{n-2}pp = CnD_{n-2}n = Cn1 = 1$. If $p = 1$, finally, $E_{n-2}pp = C1D_{n-2}1 = C1n = n$.

THEOREM 5. *For every p in L_n, $F_{n-2}p = 1$.*

PROOF. If $p \neq 1$, then, by Theorem 4, we have

$$F_{n-2}p = E_{n-2}E_{n-2}ppE_{n-2}E_{n-2}ppE_{n-2}pp = E_{n-2}1E_{n-2}11 = E_{n-2}1n = C1D_{n-2}n = C11 = 1.$$

If $p = 1$, then, again by Theorem 4,

$$F_{n-2}p = E_{n-2}E_{n-2}11E_{n-2}E_{n-2}11E_{n-2}11 = E_{n-2}nE_{n-2}nn = E_{n-2}n1CnD_{n-2}1 = Cnn = 1.$$

THEOREM 6. *For every p in L_n, $M_{n-2}p = Np$.*

PROOF. $M_{n-2}p = E_{n-2}pF_{n-2}p = E_{n-2}p1 = CpD_{n-2}1 = Cpn = Np$.

THEOREM 7. *For every p and q in L_n, $I_{n-2}pq = Cpq$.*

PROOF.

$$I_{n-2}pq = E_{n-2}pE_{n-2}F_{n-2}qq = E_{n-2}pE_{n-2}1q = E_{n-2}pC1D_{n-2}q = E_{n-2}pD_{n-2}q = CpD_{n-2}D_{n-2}q.$$

But, by Theorem 3, we have $D_{n-2}D_{n-2}q = q$. Hence $I_{n-2}pq = Cpq$.

Thus we have shown that in each L_n it is possible to define in terms of Cpq and Np a function, namely, $E_{n-2}pq$, in terms of which Cpq and Np are again definable.

NEW YORK UNIVERSITY

REMARKS ON THE INDUCTIVE PRINCIPLE AND RELATED EXISTENCE THEOREMS†

BY HENRY BLUMBERG

It is the purpose of the present note to show the validity, in a natural formulation, of the inductive principle and of related existence theorems for the general linear order and for the n-fold order.‡ The simple relations which these theorems bear to the inductive principle and to one another are brought out tersely. An example shows that the inductive principle is not valid for the general $\aleph_0$-fold order.

Let A be a given linear order. By an *extension of an initial segment* I of A—analog of neighborhood of a point—we understand a segment§ of A containing elements of both‖ I and $\bar{I}$. If E ($\neq A$) is a subset of A, let I be the set of elements of A preceding all the elements of E; we refer to I as *the initial segment associated with* E. While I consists exclusively of elements of $\bar{E}$, every extension of it contains elements of E. To deny the existence of an initial segment I thus characterized is to affirm that $\bar{E} \equiv A$. In other words, to say that for every initial segment consisting exclusively of elements of $\bar{E}$ there is an extension consisting exclusively of elements of $\bar{E}$ is to say that $\bar{E} \equiv A$. The inductive principle for a general linear order is thus a rewording of the characteristic property of the initial segment

---

† Presented to the Society, April 7, 1928.

‡ The inductive principle appears here in a different light from that of the view represented, for example, by H. Poincaré in *Science and Hypothesis*. In this connection, see also Khintchine, *Das Stetigkeitsaxiom des Linearcontinuums als Inductionsprinzip betrachtet*, Fundamenta Mathematicae, vol. 4 (1923), p. 164, and Hildebrandt, *The Borel theorem and its generalizations*, this Bulletin, vol. 32 (1926), p. 423.

§ A segment of A is a subset of A containing with every pair of its elements all the elements of A between them.

‖ We denote the complement $A - E$ of a set E by $\bar{E}$. If $I = 0$, where 0 stands for the null initial segment, we regard every non-null initial segment of A as an extension of I; if $I = A$, we regard every non-null final segment of A as an extension of I.

associated with a subset of A. We call the element property α *inductive for* A if its validity for all the elements of an initial segment I of A implies its validity for all the elements of some extension of I.† We thus have the following principle.

INDUCTIVE PRINCIPLE FOR A LINEAR ORDER. *If A is a linear order, and α an element-property inductive for A, every element of A has property α.*

If A is normally ordered, it is special in that every initial segment has a first element after it; the inductive principle thus takes the familiar form in this case. If A is the linear continuum, it is special in that every initial segment has a last element, or a first following it, and the inductive principle again reduces to known form.

As an application of the inductive principle, we may derive the Borel Covering Theorem for a general linear order. For suppose that T is a given set of segments of A such that every initial segment of A has an extension which is an element of T. Let A_1 be the linear order whose elements are the initial segments of A, it being understood that if I_1 and I_2 are two elements of A_1 such that $I_1 \subset I_2$, then $I_1 < I_2$. We say that an element of A_1 is *finitely coverable*, if it is contained in the sum of a finite number of elements of T. Let J be an initial segment of A_1 all of whose elements are finitely coverable; I the sum of the elements of J; X an element of T which is an extension of I; e an element of IX; and $I(e)$ the set of elements of A preceding e. Since e belongs to I, $I(e)$ belongs to J and is therefore finitely coverable. Hence $I+X$ is finitely coverable, and therefore every element of A_1 lying in $I+X$ is finitely coverable. There thus exists for J an extension all of whose elements are finitely coverable. Accordingly, finite coverability is an inductive property for A_1, and since A belongs to A_1, it is finitely coverable. We thus have the following theorem.

BOREL COVERING THEOREM FOR A LINEAR ORDER. *If T is a set of segments of the linear order A such that every initial segment of A has an extension which is an element of T, there exists a finite subset of T having the same property.*

† Every element property is regarded as *vacuously* valid for the elements of the null set.

The Borel theorem remains valid if we replace A by a segment S of it, provided the given set T contains an extension of every initial segment of S, with the understanding that an extension of the initial segment I of S means an extension of $J+I$, where J is the initial segment of A associated with S.

We mention lastly, as immediate implication of the Borel theorem, the existence theorem associated with a *descending segment property*, that is, a property such that if it holds for a segment S of A, and S is contained in the sum of a finite number of segments $S_1, \cdots, S_n$, it holds for at least one S_ν.

If α is a descending segment property, and S a segment of A having property α, then for some initial segment of S every extension has property α.

For otherwise every initial segment of S would have an extension not having property α, and therefore, according to the Borel theorem, and contrary to the descending character of α, S would be contained in a finite number of segments none of property α.

We next consider a general *n-fold order* $A = A_1 A_2 \cdots A_n$, the A_ν being any linear orders whatsoever, and the set A consisting of all the elements $a = (a_1, a_2, \cdots, a_n)$, where the νth *coordinate* a_ν is an element of A_ν. By an *initial rectangle* of A, we understand an n-fold order $I = I_1 I_2 \cdots I_n$, where I_ν is an initial segment of A_ν. An *extension* of I is the set of elements in an order $X = X_1 X_2 \cdots X_n$, where X_ν is an extension of I_ν. The element property α will be said to be *inductive* for A, if for every initial rectangle with elements all of property α, there exists an extension with elements all of property α. We now prove the following principle.

INDUCTIVE PRINCIPLE FOR AN n-FOLD ORDER. *If A is an n-fold order, and α an element property inductive for A, every element of A has property α.*

PROOF. Since the initial *null rectangle*, that is, the rectangle for which the component linear initial segments are all null sets, has (vacuously) all its elements of property α, there is an extension of it, that is, an initial rectangle with every compo-

nent $\neq 0$, such that every element of it has property† α. Let $I = I_1 I_2 \cdots I_n$ be any given initial rectangle such that $I_\nu = 0$, $(\nu = 1, \cdots, n)$, and every element of I has property α. For the order I_n, let J_n be the associated order of its initial segments. We shall say that the element I_n' of J_n has property β if there is an extension $X^{(n-1)} = X_1 X_2 \cdots X_{n-1}$ of $I^{(n-1)} = I_1 I_2 \cdots I_{n-1}$ such that every element of $X^{(n-1)} I_n' = X_1 X_2 \cdots X_{n-1} I_n'$ has property α. Suppose that J_n' is an initial segment of J_n every element of which has property β. Let I_n^* be the initial segment of I_n which is the sum of the elements of J_n'. Since every element of $I^{(n-1)} I_n^*$ has property α, which is inductive by hypothesis, there is an extension $X_1 \cdots X_{n-1} X_n = X^{(n-1)} X_n$ of this rectangle having only elements of property α. Let e be an element common to X_n and I_n^*. The element e belongs to an element of J_n', therefore, if $I(e)$ represents the set of elements of I_n preceding e, $I(e)$ is a subset of an element of J_n'. Therefore, on account of property β, there is an extension $X'^{(n-1)} = X_1' X_2' \cdots X_{n-1}'$ of $I^{(n-1)}$ such that every element of $X'^{(n-1)} I(e)$ has property α. If $X''^{(n-1)}$ is an extension of $I^{(n-1)}$ common to $X^{(n-1)}$ and $X'^{(n-1)}$, every element of $X''^{(n-1)} I_n'$, where $I_n' = I(e) + X_n$, has property α. Every element of J_n' contained in I_n' thus has property β. Property β is therefore inductive for J_n, and is thus valid for all of its elements, in particular for I_n. Consequently, there is an extension $Y^{(n-1)} = Y_1 Y_2 \cdots Y_{n-1}$ of $I^{(n-1)}$ such that every element of $Y^{(n-1)} I_n$ has property α. Let us now say, holding I_n fixed, that an element $e^{(n-1)} = e_1 e_2 \cdots e_{n-1}$ of $A^{(n-1)} = A_1 A_2 \cdots A_{n-1}$ has property α' if $e_1 e_2 \cdots e_{n-1} e_n$ has property α for every element e_n of I_n. We have shown that if $I^{(n-1)}$ is such that every element of $I^{(n-1)} I_n$ has property α, there is an extension $Y^{(n-1)}$ of $I^{(n-1)}$ such that every element of $Y^{(n-1)} I_n$ has

† It is to be understood that the meaning of vacuous validity is not to be so interpreted as to permit us to reason that if there is at least one component I_ν of the initial rectangle $I = I_1 I_2 \cdots I_n$ which is the null set, then α is (vacuously) valid for the elements of I, and therefore an extension of I exists with all of its elements of property α. But while such reasoning is disallowed, we are to understand that it is permissible to conclude that such an extension exists in case it is already known that α is valid for every element of I', where I' is an initial rectangle with components identical with those of I—now assumed all $\neq 0$ except for I_ν—except that the νth component of I' is not the null set. The present remark relates also to the example given below of an $\aleph_0$-fold order for which the inductive principle is not valid.

property α. In other words, property α' is inductive for the $(n-1)$-fold order $A^{(n-1)}$. Assuming the inductive principle for every $(n-1)$-fold order, we conclude that every element of $A^{(n-1)}I_n$ has property α. Similarly we conclude that every element of $A^{(n-1)}A_n = A$ has property α, that is, that α is inductive for A.

That the inductive principle need not hold for an $\aleph_0$-fold order is seen from the following example. Let $A = A_1A_2A_3 \cdots$ be the $\aleph_0$-fold order in which every A_ν is the order of the positive integers in ascending magnitude. We shall say that $e = (e_1, e_2, \cdots)$ of A has property α if $\sum_1^\infty 1/e_\nu$ is divergent. Suppose that $I = I_1I_2 \cdots$ is an initial segment of A for all the elements of which α is valid. Let μ_ν be the last element of I_ν, in case there is such a last element. Take as extension of I the $\aleph_0$-fold order $X = X_1X_2 \cdots$, where X_ν consists of the pair of elements μ_ν, $\mu_\nu+1$ when μ_ν exists, and is an arbitrarily chosen extension, that is, an arbitrary non-null final segment of I_ν, when μ_ν does not exist. Since $\sum'1/\mu_\nu$, where the summation $\sum'$ applies to the ν for which μ_ν exists, is divergent, it follows that $\sum'1/(\mu_\nu+1)$ is divergent. Therefore $\sum_1^\infty 1/e_\nu$ is divergent for every element of X, so that for every initial rectangle whose elements have property α, there is an extension whose elements have property α. Therefore α is inductive without being valid for every element of A.

This example fortifies the idea that while the *passage* from a set to its complement is *localizable* in the case of a linear order, and in a certain distinct sense also in the case of the n-fold order, it is not so in the case of the $\aleph_0$-fold order.

The Borel theorem and the theorem on the descending segment property are extensible to the n-fold order without novelty of method.

In the case of a linear order, we may, of course, substitute for the definitional property of an extension X of an initial segment I the property that X is a non-null initial segment of $A-I$. But if we define extension for an n-fold order on the basis of this modification, the inductive principle need obviously no longer hold for an n-fold order, with $n \geqq 2$, as we see from the following example: $A = A_1A_2$, A_1 and A_2 both consisting of the elements e_1 and e_2 with $e_1 \prec e_2$; property α is to be valid for (e_1, e_1) and (e_2, e_2) but not for (e_1, e_2) and (e_2, e_1).

THE OHIO STATE UNIVERSITY

NOTE ON AN ASSOCIATIVE DISTRIBUTIVE ALGEBRA IN WHICH THE COMMUTATIVE LAW OF ADDITION DOES NOT HOLD

BY H. S. VANDIVER

1. *Introduction.* We shall give here a description of an algebra in which the elements are closed under the associative laws of addition and multiplication, a left and a right distributive law holds connecting addition and multiplication, and the usual equality axioms hold, yet the commutative law of addition does not hold.

2. *Certain Matrices.* As in other papers,* we define a semi-group as a set of elements closed under an associative operation, and a semi-ring as an algebra of the type above described. In the papers just cited we considered a set of elements

$$C_1, C_2, \cdots, C_i, \cdots, C_j; \qquad (j \geqq i), \tag{1}$$

with $C_{i+1} = C_i$ and forming a semi-ring, with

$$C_m + C_n = C_{m+n}, \qquad C_m \cdot C_n = C_{mn}.$$

On page 581 of the first paper mentioned, we showed how to adjoin a zero element to this set, say C_0.

In (1), set $i=j=1$, and consider the only two distinct elements in the set, after adjoining C_0. We then have the relations

$$\begin{gathered} C_1 + C_1 = C_1; \quad C_1 + C_0 = C_0 + C_1 = C_1; \quad C_0 + C_0 = C_0; \\ C_1 \cdot C_1 = C_1; \quad C_0 \cdot C_1 = C_0 \cdot C_1 = C_0. \end{gathered} \tag{2}$$

Note that these elements are *not* isomorphic to the two residue classes modulo 2, and the cancellation law of addition does not hold in general; that is, from $C_1 + C_0 = C_1 + C_1$, we cannot infer $C_0 = C_1$. Consider the matrices

$$A = \begin{pmatrix} C_0 & C_1 \\ C_0 & C_1 \end{pmatrix}; \qquad B = \begin{pmatrix} C_1 & C_0 \\ C_1 & C_0 \end{pmatrix}; \qquad C = \begin{pmatrix} C_1 & C_1 \\ C_1 & C_1 \end{pmatrix}.$$

* Proceedings of the National Academy of Sciences, vol. 20 (1934), p. 581; vol. 21 (1935), p. 162. This Bulletin, vol. 40 (1934), pp. 916–920.

Employing the usual definitions of addition and multiplication of matrices and (2), we have

$$(3)\qquad \begin{aligned} &A+B=C;\quad A+A=A;\quad B+B=B;\\ &C+C=C;\quad A+C=C;\quad B+C=C. \end{aligned}$$

$$(4)\qquad \begin{aligned} &A^2=A;\quad B^2=B;\quad C^2=C;\quad AB=B;\quad AC=C;\\ &BC=C;\quad BA=A;\quad CA=A;\quad CB=B. \end{aligned}$$

We define a *left identity* in a semi-group S as an element E_l such that $E_lM=M$ for any M in S, with an analogous definition of *right identity*. An element O_r such that $MO_r=O_r$ for each M in S, will be called a *right annulator* (or annihilator) in S, and we define *left annulator* in an analogous way. If an element is both a right and left annulator it is called an *annulator*. We note from (3) that A, B, and C form a semi-group under addition in which C is an annulator. They form a semi-group under multiplication in which each element is a right-hand annulator and a left-hand identity. By known properties of matrices, since C_0 and C_1 form a semi-ring, then A, B, and C generate a semi-ring. This result follows also from (3) and (4).

3. *Description of the System.* Consider a set of bracket symbols $[\alpha, \beta]$, where α and β are any of the matrices A, B, or C. We say that $[\alpha, \beta]=[\gamma, \delta]$ if and only if $\alpha=\gamma$, $\beta=\delta$. Hence there are nine distinct symbols. Also we define addition and multiplication as follows:

$$\begin{aligned} [\alpha, \beta]+[\gamma, \delta] &= [\alpha\gamma, \beta+\delta],\\ [\alpha, \beta]\cdot[\gamma, \delta] &= [\alpha\gamma, \beta\delta]. \end{aligned}$$

From these definitions we note that each element is a right annulator and a left identity under multiplication. Hence the associative law of multiplication and the right and left distributive laws are easily seen to hold. The associative law of addition holds, since

$$\begin{aligned} ([\alpha, \beta]+[\gamma, \delta])+[\epsilon, \theta] &= [\alpha\gamma, \beta+\delta]+[\epsilon, \theta]\\ &= [\alpha\gamma\epsilon, (\beta+\delta)+\theta] = [\alpha(\gamma\epsilon), \beta+(\delta+\theta)]\\ &= [\alpha, \beta]+([\gamma, \delta]+[\epsilon, \theta]). \end{aligned}$$

Hence the bracket symbols form a semi-ring. The commutative law of addition does not hold in general in this semi-ring since

$$[A, \beta] + [B, \gamma] = [AB, \beta + \gamma],$$
$$[B, \gamma] + [A, \beta] = [BA, \gamma + \beta],$$

but $BA \neq AB$. These symbols have a property under addition which might be called quasi-commutativity:

$$[\alpha, \beta] + [\alpha, \beta] + [\gamma, \delta] + [\gamma, \delta] = [\alpha, \beta] + [\gamma, \delta] + [\alpha, \beta] + [\gamma, \delta],$$

for the left-hand member reduces to $[\gamma, \beta+\beta+\delta+\delta]$ and the right to $[\gamma, \beta+\delta+\beta+\delta]$, which are equal since A, B, and C are commutative under addition. It is also easy to see that $MNMN = MMNN$, for M and N are bracket symbols.

THE UNIVERSITY OF TEXAS

BRANCHED AND FOLDED COVERINGS*

BY A. W. TUCKER

A simple example of a *branched* covering arises when one sphere is mapped on another so that each point of the first sphere goes into the point of the second which has the same latitude but double the longitude. This is a covering of degree two with simple branching at the north and south poles. As an example of a *folded* covering we take a torus, thought of as a sphere with a handle on one side, and project it radially inward on a smaller concentric sphere. The torus covers the sphere once but with a fold produced by collapse of the handle. The product of this torus-sphere covering with the previous sphere-sphere covering yields a torus-sphere covering of degree two which is both *branched and folded*. Suitable triangulation of the torus and the spheres will turn the above mappings into simplicial mappings in which each simplex maps barycentrically into a simplex of the same dimension. In what follows we make some rudimentary calculations concerning the branching and folding of a simplicial covering of one n-dimensional complex by

* Presented to the Society, September 1, 1936.

another. We obtain two formulas which connect the Euler characteristics (numbers) of the complexes with those of the subcomplexes about which the branching and folding occur.

Let $K \to K'$ be a (continuous) simplicial mapping of one simplicial n-complex, K, on another, K'. For the sake of simplicity we assume that K and K' are absolute orientable manifolds* and that each p-simplex s of K goes into a p-simplex s' of K', that is, that no simplex s collapses into one of lower dimension. Later, however, we shall weaken these assumptions in various ways. Let the fundamental n-cycles C, C' of K, K' be oriented so that $C \to dC'$, where the integer $d \geqq 0$; d is the *degree* of the mapping. Let $C(s)$ denote the part of C which lies on the star of the simplex s, and $C'(s')$ the part of C' which lies on the star of the corresponding s'. Then the integer $d(s)$ defined by $C(s) \to d(s)C'(s')$ measures the degree of s in the mapping. Normally $d(s) = 1$; if $d(s) \neq 1$ we say that s is *exceptional* in the mapping. If $d(s) > 1$ the *exceptionality* is a branching about s of multiplicity $d(s) - 1$. If $d(s) = 0$, s belongs to the crease of a fold in the covering. If $d(s) < 0$, s belongs to a part of K which gives a negatively sensed layer of a fold; s contributes simply to the layer if $d(s) = -1$, otherwise there is branching about s of multiplicity $|d(s)| - 1$. Our use of the terms *branching* and *fold* is in accord with their customary meaning as applied to coverings, but for the purpose of this paper the descriptions just given may be regarded as definitions of these terms.

Let the integer $e(s) = d(s) - 1$ be taken as a measure of the *exceptionality* of s in the mapping. For each integer $e \times 0$ we form the simplexes of exceptionality $e(s) = e$, if any, into one or more *subcomplexes* $K^{[e]i}$, $(i = 1, 2, \cdots, n_e)$.† In practice the $K^{[e]i}$ will arise as natural connected units about which the branching and folding (and negative stratification) occur. For example, in the torus-sphere covering of degree two described in the opening paragraph the $K^{[1]i}$ = the two branch points at the poles, the $K^{[-1]i}$ = four open segments which constitute the lines of fold, and the $K^{[-2]i}$ = an open region on the inner surface of the

* See Lefschetz, *Topology*, for the terminology.

† Subcomplex is used in the general sense of the author's thesis, *An abstract approach to manifolds*, Annals of Mathematics, (2), vol. 34 (1933), pp. 191–243, as applying to any subset of simplexes on which the (relative) boundary of a (relative) boundary is zero.

handle—a 2-cell bounded by the $K^{[-1]i}$. However, for the purpose of our calculations it does not matter just how the $K^{[e]i}$ are constructed. At worst each $K^{[e]i}$ may consist of an individual simplex.

Summing over all s which map into a given s' we have $\sum d(s) = d$, the degree of the mapping. Therefore, extending the sum over all p-simplexes s, we get $\sum d(s) = d\alpha_p'$, where α_p' is the number of p-simplexes in K'. Hence $\alpha_p + \sum e(s) = d\alpha_p'$. But $\sum e(s) = \sum_e \sum_i e\alpha_p^{[e]i}$. Therefore $\alpha_p + \sum_e \sum_i e\alpha_p^{[e]i} = d\alpha_p'$, and so

$$X + \sum_e \sum_i eX^{[e]i} = dX',$$

where $X = \sum_p (-1)^p \alpha_p, \cdots$, are the Euler characteristics of K, $\cdots$. This is the first of our two formulas. If the covering has neither branching nor folding the sum on the left side is zero and we have the familiar formula $X = dX'$. If there is branching but no folding only positive values of e occur on the left; we have a generalization of the formula used to characterize a Riemann surface by its branch-points and number of sheets. If there is folding without branching the sum on the left side is zero except for $e = -1, -2$. Values of $e < -2$ go with branching inside negative stratification. In the torus-sphere covering of degree two described in our opening paragraph $X = 0$, $X' = 2$, $d = 2$, $\sum_i (1)X^{[1]i} = 2$, $\sum_i (-1)X^{[-1]i} = 4$, $\sum_i (-2)X^{[-2]i} = -2$.

In contrast with the preceding work where orientation has played such an essential part we now turn to calculations of an absolute nature, based on $|d(s)|$ rather than on $d(s)$. One justification of this is the fact that the sign of $d(s)$ has no meaning from a local point of view—it was determined by the orientation of C, C' so that $C \to dC'$ where $d \geqq 0$. We use $\epsilon(s) = |d(s)| - 1$ as a measure of the *absolute exceptionality* of s; in terms of $\epsilon(s)$ we form subcomplexes $K^{[\epsilon]j}$, $(j = 1, 2, \cdots, n_\epsilon)$, just as we did with $e(s)$ above. Let $\delta(s')$ denote the sum $\sum |d(s)|$ taken over all s which map into a given s'; the values of $\delta(s')$ differ for different s' but are all congruent mod 2 since extra layers occur in pairs. We choose a non-negative number $\delta \leqq$ each $\delta(s')$ but congruent to each mod 2, and we set $\delta(s') = \delta + 2\lambda(s')$. The non-negative integer $\lambda(s')$ measures the number of pairs of layers there are over s' in excess of the basic number which we have chosen. From the simplexes s' for which $\lambda(s') \geqq \mu > 0$ we form

subcomplexes $K'^{[\mu]k}$, $(k=1, 2, \cdots, n_\mu)$. Then summing over all p-simplexes s we get $\sum|d(s)| = \delta\alpha_p' + 2\sum_\mu\sum_k\alpha_p'^{[\mu]k}$. Hence $\alpha_p + \sum\epsilon(s) = \delta\alpha_p' + 2\sum_\mu\sum_k\alpha_p'^{[\mu]k}$. But $\sum\epsilon(s) = \sum_\epsilon\sum_j\epsilon\alpha_p^{[\epsilon]j}$. Therefore $\alpha_p + \sum_\epsilon\sum_j\epsilon\alpha_p^{[\epsilon]j} = \delta\alpha_p' + 2\sum_\mu\sum_k\alpha_p'^{[\mu]k}$, and so

$$X + \sum_\epsilon\sum_j\epsilon X^{[\epsilon]j} = \delta X' + 2\sum_\mu\sum_k X'^{[\mu]k}.$$

This is the second of our two formulas. It applies even if K' is not orientable (in which case K may or may not be orientable). If K' is orientable the degree d will probably be taken as the value of δ, but this is not necessary.

We started with the assumption that K, K' were orientable absolute n-manifolds, but the only use we have made of this assumption has been for the n-cycles C, C' and the relative n-cycles $C(s)$, $C'(s')$. All we really need to assume is that K, K' are orientable n-circuits and that the star of each simplex of K' carries an irreducible basis $C'(s') \neq 0$; for the second formula we may dispense with the orientability of K' and leave the orientability of K an open question.

We may also weaken the assumption that in the mapping $K \to K'$ each simplex s goes into a simplex s' of like dimension. It is sufficient to suppose that K can be divided into p-cell-like* subcomplexes S which map into p-simplexes s'. These S take the place of the simplexes s in the preceding work; a p-cell-like S, like a p-simplex s, has an Euler characteristic $(-1)^p$. The *star* of an S would be the minimal open subcomplex containing S and composed of S's. This extension to more general simplicial coverings has importance for simplicial approximation of continuous coverings.

PRINCETON UNIVERSITY

* See Tucker, loc. cit. The following is a simple example of a 1-cell-like S mapping into a 1-simplex s'. Let three tetrahedra $ABCD$, $ABDE$, $ABEF$ be collapsed by the mapping $A \to A'$ and B, C, D, E, $F \to B'$. The three tetrahedra and all their faces and edges incident with A are mapped into $A'B'$. The aggregate of these simplexes is readily seen to be a 1-cell-like subcomplex (=a solid cone with vertex and base removed).

NOTE ON FORMULAS FOR THE NUMBER OF REPRESENTATIONS OF AN INTEGER AS A SUM OF $2h$ SQUARES*

BY R. D. JAMES

1. *Introduction.* It is known that the number of representations of an arbitrary integer n as a sum of $2h$ squares may be expressed in the form

$$\lambda \sum (-1)^{(hd-h)/2} d^{h-1}, \tag{1}$$

when $1 \leqq h \leqq 4$. The number λ may depend on the linear form of n and the summation is over all *odd* divisors d of n. For example, an odd integer can be represented as a sum of four squares in $8\sum d$ ways, while an even integer can be represented in $24\sum d$ ways. The question arises whether formulas of the type (1) hold when $h \geqq 5$. This question has been answered in the negative for $n = m$ or $n = 2m$, where m is odd, by E. T. Bell† and in some cases for $n = 2^{\alpha}m$, $\alpha \geqq 1$ by the author.‡ These results were obtained by the use of the theory of elliptic functions.

In this note we shall prove a necessary condition which does not require any elliptic function theory and which enables us to decide if a formula of the type (1) is possible or not. It is to be noticed that this method does not show when formulas of the type (1) are true but only when they are *not* true.

2. *Necessary Condition for Formulas of Type* (1). Let n_j denote any one of a given class $K(n_j)$ of positive integers in which $n_j < n_{j+1}$. For example, we might take $n_j = 4j-3$ and then $K(n_j)$ would be the class of all positive integers which are $\equiv 1 \pmod 4$. Let $N(n, 2h)$ denote the number of representations of an integer n as a sum of $2h$ squares, both the arrangement of the squares and the signs of their square roots being relevant in counting the representations. Then we have the following result, which

* Presented to the Society, April 11, 1936.

† Journal für die reine und angewandte Mathematik, vol. 163 (1930), pp. 65–70; Journal of the London Mathematical Society, vol. 4 (1929), pp. 279–285.

‡ American Journal of Mathematics, vol. 58 (1936), pp. 536–544.

is almost trivial, but is nevertheless sufficient to show the impossibility of formulas like (1) in a great number of cases.

THEOREM 1. *In order that the formula*

$$N(n, 2h) = \lambda \sum_{d|n} (-1)^{(hd-h)/2} d^{h-1} \tag{2}$$

should hold for every integer in the class $K(n_j)$ with λ independent of n, it is necessary that

$$N(n_2, 2h) \sum_{d|n_1} (-1)^{(hd-h)/2} d^{h-1} = N(n_1, 2h) \sum_{d|n_2} (-1)^{(hd-h)/2} d^{h-1}. \tag{3}$$

The summations in each case are over all odd divisors of n, n_1, and n_2, respectively.

PROOF. If the formula (2) is to hold for all n in $K(n_j)$, it must certainly hold when $n = n_1$ and $n = n_2$. The necessary condition (3) is obtained by equating the values of λ found for these two values of n.

3. *Impossibility of Formulas of Type* (2). We now give a few results to illustrate the use of Theorem 1.

THEOREM 2.* *If $h \geqq 5$ no formula of the type* (2) *can hold for all odd integers $m \equiv 1 \pmod 4$ with λ independent of m.*

PROOF. Here we have $K(n_j) = K(4j-3)$ and $n_1 = 1$, $n_2 = 5$. The condition (3) becomes

$$N(5, 2h) = N(1, 2h)[1 + 5^{h-1}]. \tag{4}$$

Since

$$1 = (\pm 1)^2 + 0^2 + \cdots + 0^2,$$

$$5 = (\pm 1)^2 + (\pm 2)^2 + 0^2 + \cdots + 0^2,$$

$$5 = (\pm 1)^2 + \cdots + (\pm 1)^2 + 0^2 + \cdots + 0^2,$$

we find that†

---

* E. T. Bell, loc. cit.

† $C_{n,t}$ denotes the binomial coefficient $n!/t!(n-t)!$

$$N(1, 2h) = 2C_{2h,1},$$
$$N(5, 2h) = 2^2C_{2h,1} \cdot C_{2h-1,1} + 2^5C_{2h,5}.$$

Substitution of these values in (4) reduces the condition to the form

$$30(2h - 1) + 2(2h - 1)(2h - 2)(2h - 3)(2h - 4) = 15(1+5^{h-1}).$$

For $h=2r+1$ this equation is identical with the one given by E. T. Bell in Theorem α of his Journal für Mathematik paper referred to above. It can be verified that this equation is true only when $1 \leqq h \leqq 4$. In fact, for $h \geqq 5$ the left side is less than the right side.

THEOREM 3.* *If h is odd and $\geqq 3$, or if h is even and $\geqq 8$, no formulas of the type* (2) *can hold for all integers of the form $2m$, m odd, with λ independent of m.*

PROOF. In this case we have $K(n_i)=K(4j-2)$ and $n_1=2$, $n_2=6$. Then (3) becomes

$$N(6, 2h) = N(2, 2h)[1 + (-1)^h 3^{h-1}].$$

Since

$$2 = (\pm 1)^2 + (\pm 1)^2 + 0^2 + \cdots + 0^2,$$
$$6 = (\pm 1)^2 + (\pm 1)^2 + (\pm 2)^2 + 0^2 + \cdots + 0^2,$$
$$6 = (\pm 1)^2 + \cdots + (\pm 1)^2 + 0^2 + \cdots + 0^2,$$

we have

$$N(2, 2h) = 2^2C_{2h,2},$$
$$N(6, 2h) = 2^3C_{2h,2} \cdot C_{2h-2,1} + 2^6C_{2h,6}.$$

After simplification the condition becomes

$$(4h - 4)[45 + (2h - 3)(2h - 4)(2h - 5)] = 45[1 + (-1)^h 3^{h-1}].$$

This equation is true only when $h=1$, 2, 4, or 6. For h odd and $\geqq 3$ the left side is greater than the right side. For h even and $\geqq 8$ the left side is less than the right side.

---

* E. T. Bell, loc. cit.

THEOREM 4. *If $h=5$ or $h\geqq 7$, no formulas of the type* (2) *can hold for all integers of the form* $2m$, $m\equiv 1 (mod\ 4)$, *with* λ *independent of* m.

PROOF. This time we have $K(n_j)=K(8j-6)$, $n_1=2$, $n_2=10$. Then (3) becomes

$$N(10, 2h) = N(2, 2h)[1 + 5^{h-1}].$$

Since

$$10 = (\pm 1)^2 + (\pm 3)^2 + 0^2 + \cdots + 0^2,$$
$$10 = (\pm 1)^2 + (\pm 1)^2 + (\pm 2)^2 + (\pm 2)^2 + 0^2 + \cdots + 0^2,$$
$$10 = (\pm 1)^2 + \cdots + (\pm 1)^2 + (\pm 2)^2 + 0^2 + \cdots + 0^2,$$
$$10 = (\pm 1)^2 + \cdots + (\pm 1)^2 + 0^2 + \cdots + 0^2,$$

we have

$$N(10, 2h) = 2^2C_{2,1}\cdot C_{2h,2} + 2^4C_{2h,2}\cdot C_{2h-2,2} + 2^7C_{2h,6}\cdot C_{2h-6,1} + 2^{10}C_{2h,10}.$$

After simplification the condition reduces to

$$28350[1 + (2h-2)(2h-3)] + 1260(2h-2)\cdots(2h-6) + 2(2h-2)\cdots(2h-9) = 14175[1 + 5^{h-1}].$$

It can be verified that this equation is true for $1\leqq h\leqq 6$, $h\neq 5$, but that the left side is less than the right side when $h\geqq 7$, and the right side is less than the left side when $h=5$.

THE UNIVERSITY OF CALIFORNIA

ON HIGHER DERIVATIVES OF ORTHOGONAL POLYNOMIALS*

BY H. L. KRALL

1. *Introduction.* Let $\{\phi_n(x)\}$ be a set of orthogonal polynomials in a finite interval (a, b) with the integrable (L) weight function† $p(x)$, that is,

$$\int_a^b p(x)\phi_n(x)\phi_m(x)dx = 0, \qquad (n \neq m),$$

$$p(x) \geqq 0, \quad \int_a^b p(x)dx > 0, \quad \phi_n(x) = x^n + a_{n,n-1}x^{n-1} + \cdots + a_{n0}.$$

It has been shown‡ that if the first derivatives $\{\phi_n'(x)\}$ also form a set of orthogonal polynomials, then the original set are Jacobi polynomials. The purpose here is to show that if the rth derivatives $\{\phi_n^r(x)\}$ form an orthogonal set, then again $\{\phi_n(x)\}$ is a set of Jacobi polynomials. The proof is based on the following lemma.§

LEMMA. *Let $Q(x)$ be non-negative in the (finite or infinite) interval (c, d), and such that the constants β defined by the formula*

$$\beta_k = \int_c^d Q(x)x^k dx, \qquad (k = 0, 1, \cdots),$$

exist, and for a certain positive integer r

$$\int_c^d Q(x)\phi_n(x)G_{n-r-1}(x)dx = 0, \qquad (n = r+1, r+2, \cdots),$$

---

* Presented to the Society, April 11, 1936.

† There is no restriction in assuming (as we do) that if α, β are any two numbers, $a<\alpha<b<\beta$, then $p(x)\neq 0$ almost everywhere in (a, α); $p(x)\neq 0$, almost everywhere in (β, b).

‡ W. Hahn, *Über die Jacobischen Polynome und zwei verwandte Polynomklassen*, Mathematische Zeitschrift, vol. 39 (1935), pp. 634–638. H. L. Krall, *On derivatives of orthogonal polynomials*, this Bulletin, vol. 42 (1936), pp. 423–428.

§ See Krall, loc. cit.

where $G_n(x)$ is an arbitrary polynomial of degree $\leqq n$. Then almost everywhere

$$Q(x) = \begin{cases} P_r(x)p(x) \text{ in } (a, b), \\ 0 \text{ elsewhere}, \end{cases}$$

where $P_r(x)$ is a polynomial of degree $\leqq r$.

2. *Identity of the Intervals of Orthogonality.* The $\{\phi_n(x)\}$ satisfy the recurrence relations

$$\phi_{n+2}(x) = (x - c_{n+2})\phi_{n+1}(x) - \lambda_{n+2}\phi_n(x), \quad (c_{n+2}, \lambda_{n+2}, \text{constants}).$$

Differentiating this r times, we obtain

$$\begin{aligned} \phi'_{n+2}(x) &= (x - c_{n+2})\phi'_{n+1}(x) - \lambda_{n+2}\phi'_n(x) + \phi_{n+1}(x), \\ \phi''_{n+2}(x) &= (x - c_{n+2})\phi''_{n+1}(x) - \lambda_{n+2}\phi''_n(x) + 2\phi'_{n+1}(x), \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \phi^{r}_{n+2}(x) &= (x - c_{n+2})\phi^{r}_{n+1}(x) - \lambda_{n+2}\phi^{r}_n(x) + r\phi^{r-1}_{n+1}(x). \end{aligned} \tag{1}$$

Let $q(x)$ be the weight function of the orthogonal set $\{\phi_n{}^r(x)\}$ in the interval (c, d). If we multiply the last equation of (1) by $q(x)G_{n-r-1}(x)$ and integrate, we get

$$\int_c^d q(x)\phi^{r-1}_{n+1}(x)G_{n-r-1}(x)dx = 0, \text{ or } \int_c^d q(x)\phi^{r-1}_n(x)G_{n-r-2}(x)dx = 0.$$

In this way we obtain successively

$$\begin{aligned} &\int_c^d q(x)\phi^{r-1}_n(x)G_{n-r-2}(x)dx = 0, \quad \int_c^d q(x)\phi^{r-2}_n(x)G_{n-r-3}(x)dx = 0, \\ &\cdots, \quad \int_c^d q(x)\phi_n(x)G_{n-2r-1}(x)dx = 0. \end{aligned} \tag{2}$$

The lemma can be applied to the last equation, whence

$$q(x) = P_{2r}(x)p(x), \qquad (a, b) \equiv (c, d).$$

3. *Existence of $q^r(x)$.* Consider the function

$$S(x) = \int_a^x dx_1 \int_a^{x_1} dx_2 \cdots \int_a^{x_{r-1}} (f_r t^r + f_{r-1}t^{r-1} + \cdots + f_0)p(t)dt$$

$$= \frac{1}{(r-1)!}\int_a^x (x-t)^{r-1}(f_r t^r + \cdots + f_0)p(t)dt,$$

where the constants $f_r, f_{r-1}, \cdots, f_0$ are determined so that

$$S(b) = S'(b) = \cdots = S^{r-1}(b) = 0, \qquad \int_a^b S(x)dx = \int_a^b q(x)dx.$$

From the definition of $S(x)$, it follows that $S(a) = S'(a) = \cdots = S^{r-1}(a) = 0$, and $S^r(x) = (f_r x^r + \cdots + f_0)p(x)$. Successive integration by parts gives

$$\begin{aligned}(3)\qquad \int_a^b S(x)\phi_n^r(x)dx &= \int_a^b S'(x)\phi_n^{r-1}(x)dx = \cdots \\ &= \int_a^b S^r(x)\phi_n(x)dx = 0, \qquad (n \geqq r+1).\end{aligned}$$

Thus

$$(4)\qquad \int_a^b S(x)\phi_n^r(x)dx = \int_a^b q(x)\phi_n^r(x)dx = 0, \qquad (n \geqq r+1).$$

Since $\int_a^b S(x)dx = \int_a^b q(x)dx$, and $\phi_n{}^r(x)$ is of degree $n-r$, we have

$$\int_a^b x^n S(x)dx = \int_a^b x^n q(x)dx, \qquad (n \geqq 0),$$

so that $q(x) \equiv S(x)$ almost everywhere. Thus $q^r(x)$ exists almost everywhere and

$$\begin{aligned}q^r(x) &= P_r(x)p(x), \\ (5)\qquad q(a) &= q'(a) = \cdots = q^{r-1}(a) = q(b) \\ &= q'(b) = \cdots = q^{r-1}(b) = 0.\end{aligned}$$

4. *Relations for the Derivatives of* $q(x)$. An integration by parts applied to the next to last equation of (2) gives

$$\begin{aligned}0 = \int_a^b q(x)\phi_n'(x)G_{n-2r}(x)dx &= \int_a^b q'(x)G_{n-2r}(x)\phi_n(x)dx \\ &+ \int_a^b q(x)G_{n-2r}'(x)\phi_n(x)dx.\end{aligned}$$

Since $q(x)=P_{2r}(x)p(x)$, the second integral is zero. The other equations of (2) give us successively

$$\int_a^b q'(x)G_{n-2r}(x)\phi_n(x)dx = 0, \quad \int_a^b q''(x)G_{n-2r+1}(x)\phi_n(x)dx = 0,$$

$$\cdots, \quad \int_a^b q^{r-1}(x)G_{n-r-2}(x)\phi_n(x)dx = 0.$$

The lemma applies to these equations and we have

$$(6) \qquad \begin{aligned} q(x) &= P_{2r}(x)p(x), & q'(x) &= P_{2r-1}(x)p(x), \cdots, \\ q^{r-1}(x) &= P_{r+1}(x)p(x), & q^r(x) &= P_r(x)p(x). \end{aligned}$$

5. *Determination of* $p(x)$. If we eliminate $p(x)$ from the last two equations of (6), we get

$$(7) \qquad q^r(x) = \frac{P_r(x)}{P_{r+1}(x)}q^{r-1}(x), \qquad q^{r-1}(a) = q^{r-1}(b) = 0.$$

We conclude that $P_{r+1}(a)=0$. For suppose not, then $q^r(a)=0$, and (7) is a linear differential equation for $q^{r-1}(x)$, with coefficients analytic at $x=a$. The initial conditions make $q^{r-1}(x)\equiv 0$ (and therefore $p(x)\equiv 0$) in some neighborhood of $x=a$. But this contradicts the condition of a preceding footnote. Similarly, $P_{r+1}(b)=0$.

The same type of argument shows that for $n=1, 2, \cdots, r$, $P_{r+n}(x)$ has n-fold zeros at both $x=a$ and $x=b$. Hence

$$P_{2r}(x) = k(x-a)^r(x-b)^r,$$

$$P_{2r-1}(x) = c(x-a)^{r-1}(x-b)^{r-1}(lx+m).$$

The first two equations of (6) now yield

$$\frac{q'(x)}{q(x)} = \frac{c(x-a)^{r-1}(x-b)^{r-1}(lx+m)}{k(x-a)^r(x-b)^r} = \frac{\alpha}{x-a} - \frac{\beta}{b-x},$$

$$q(x) = K(x-a)^\alpha(b-x)^\beta, \quad p(x) = C(x-a)^{\alpha-r}(b-x)^{\beta-r}.$$

Since this is the weight function of Jacobi polynomials, our proposition is proved.

PENNSYLVANIA STATE COLLEGE

ON ARITHMETIC INVARIANTS OF BINARY CUBIC AND BINARY TRILINEAR FORMS*

BY RUFUS OLDENBURGER

1. *Introduction.* It is the purpose of this note to show that binary cubic and binary trilinear forms can be completely and very simply characterized by *arithmetic rank invariants* for non-singular linear transformations in the complex field. We define the factorization rank† of a matrix $A=(a_{ijk})$ of order n and its associated trilinear form to be the minimum value of ϵ such that A can be "factored" into the form

$$(1) \qquad A=\left(\sum_{\alpha=1}^{\epsilon} a_{\alpha i}b_{\alpha j}c_{\alpha k}\right), \qquad (i, j, k=1, 2, \cdots, n).$$

Hitchcock‡ obtained minimum values of ϵ for certain polyadics. The number ϵ is invariant under non-singular linear transformations on the variables in the trilinear form

$$\sum a_{ijk}x_iy_jz_k$$

associated with A. In a paper which appeared in this Bulletin,§ I classified binary trilinear forms by means of ranks and a property of invariant factors, which are invariant under non-singular linear transformations in the complex field. *The ranks of that paper and the rank defined above form a complete invariant system for these forms. For binary cubic forms the factorization rank alone forms a complete invariant system.*

2. *Factorization Ranks.* The canonical binary trilinear forms are

---

* Presented to the Society, April 11, 1936.

† This rank was used in developing a general theory of non-singular p-way matrices in the paper *Non-singular multilinear forms and certain p-way matrix factorizations*, Transactions of this Society, vol. 39 (1936), pp. 422–455. If factorization rank is similarly defined for 2-way matrices, it is found that the factorization rank of a 2-way matrix is n if and only if its ordinary rank is n.

‡ F. L. Hitchcock, *The expression of a tensor or polyadic as a sum of products*, Journal of Mathematics and Physics, Massachusetts Institute of Technology, vol. 6 (1927), pp. 164–189.

§ *On canonical binary trilinear forms*, vol. 38 (1932), pp. 385–387.

$$R = x_1y_1z_1 + x_2y_2z_2,$$
$$L = x_1y_1z_1 + x_2y_2z_1 + x_2y_1z_2,$$
$$H = x_1y_1z_1 + x_1y_2z_2,$$
$$K = x_1y_1z_1.$$

The forms H, K can be distinguished from each other and from L, R by the ranks of my Bulletin paper referred to above. Since the i-, j-, and k-ranks* of L equal 2, the factorization rank ϵ of L is $\geqq 2$. If ϵ were equal to 2 for L, the form L would be equivalent to R, which is impossible. For then the matrix A of L is of the form (1), where $A=(a_{\alpha i})$, $B=(b_{\alpha j})$, $C=(c_{\alpha k})$ are non-singular, whence transformations on L with A^{-1}, B^{-1}, C^{-1} give R. Since the matrix of L can be written in the form (1), where

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{pmatrix}, \qquad B = \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad C = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \end{pmatrix},$$

it follows that $\epsilon=3$.

The factorization rank ϵ for H is $\geqq 2$, since the j- and k-ranks of H equal 2. That $\epsilon=2$ is evident since the matrix of H can be written in the form (1), where

$$A = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}, \qquad B = C = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

The matrix of K can be written in the form (1), where

$$A = B = C = (1, 0),$$

whence $\epsilon=1$.

The canonical binary cubic forms for the complex field are well known.† They are

$$P = x^3 + y^3, \qquad Q = x^2y, \qquad S = x^3.$$

Since the matrix of P can be taken to be the same as the matrix of R, its factorization rank is 2. The matrix of $3Q$ can be taken to be the matrix A whose only non-vanishing elements are

* The i-rank of $A=(a_{ijk})$, $(i, j, k=1, 2)$, is 2 if the minors (a_{1jk}) and (a_{2jk}) are linearly independent.

† L. E. Dickson, *Modern Algebraic Theories*, pp. 7–9.

$a_{112}=a_{121}=a_{212}=1$. The trilinear form associated with this matrix is $x_1y_1z_2+x_1y_2z_1+x_2y_1z_1$, which is equivalent to L under the transformations $x_1=x_2'$, $x_2=x_1'$. Since the matrices of S and K can be taken to be the same, the factorization rank of S is 1. We have proved the following result.

THEOREM. *The factorization ranks of the forms in the sets* (K, S), (R, P, H), *and* (L, Q) *are* 1, 2, *and* 3, *respectively.*

The equivalence of cubics to P, Q, S can be recognized very simply without the use of factorization rank from the theory of my previous Bulletin paper.

ARMOUR INSTITUTE OF TECHNOLOGY

A NOTE ON THE DEGREE OF POLYNOMIAL APPROXIMATION*

BY J. H. CURTISS

Let C be a rectifiable Jordan curve of the finite z plane. We shall say that a function $f(z)$ belongs to the class Lip (C, j, α) if $f(z)$ is regular in the limited region bounded by C (which we shall call the interior of C), if $f(z)$ is continuous in the corresponding closed region, and if the jth derivative of $f(z)$ is also continuous in this closed region and satisfies a Lipschitz condition with exponent α on C:

$$\left| f^{(j)}(z_1) - f^{(j)}(z_2) \right| \leqq M \left| z_1 - z_2 \right|^{\alpha},$$

z_1, z_2 on C. The number α will be positive and not greater than unity. The number j will be a positive integer or zero; we define $f^{(0)}(z)$ to be identically $f(z)$. The object of this note is to establish the following existence theorem.

THEOREM. *Let the point set* S *consist of a finite number of closed limited Jordan regions of the* z *plane bounded by the mutually exterior analytic curves* $C_1, C_2, \cdots, C_\lambda$. *Let the functions* $f_1(z), f_2(z), \cdots, f_\lambda(z)$ *belong respectively to the classes*

$$\text{Lip } (C_1, j, \alpha), \quad \text{Lip } (C_2, j, \alpha), \quad \cdots, \quad \text{Lip } (C_\lambda, j, \alpha),$$

---

* Presented to the Society, December 31, 1935.

and let $f(z) \equiv f_\nu(z)$, *where* z *is on and interior to* C_ν, *for* $\nu = 1, 2, \cdots, \lambda$. *Then there exist polynomials* $p_n(z)$ *of respective degrees* n,* $(n = 1, 2, \cdots)$, *such that* $f(z) - p_n(z) = O(n^{-j-\alpha})$ *uniformly for* z *on* S.

The theorem for the case in which S is the closed interior of the circle γ: $|z| = 1$ is an immediate consequence of certain results of Bernstein and Jackson on approximation by trigonometric sums. For let the function $f(z)$ belong to the class Lip $(\gamma, 0, \alpha)$. If we write $z = re^{i\theta}$ and $f(e^{i\theta}) \equiv u(\theta) + iv(\theta)$, it is easily verified that both $u(\theta)$ and $v(\theta)$ satisfy Lipschitz conditions with exponent α. The work of Bernstein and Jackson† proves that there exist two trigonometric sums of the nth order,

$$U_n(\theta) = \frac{1}{2} d_{n,0} a_0 + \sum_{k=1}^{n} d_{n,k}(a_k \cos k\theta + b_k \sin k\theta),$$

$$V_n(\theta) = \frac{1}{2} d_{n,0} g_0 + \sum_{k=1}^{n} d_{n,k}(g_k \cos k\theta + h_k \sin k\theta),$$

such that uniformly for all values of θ, $u(\theta) - U_n(\theta) = O(n^{-\alpha})$ and $v(\theta) - V_n(\theta) = O(n^{-\alpha})$. The numbers a_k and b_k are the Fourier coefficients of the function $u(\theta)$ and the numbers g_k and h_k are the Fourier coefficients of the function $v(\theta)$. Let $f(z) = \sum_{k=0}^{\infty} A_k z^k$, $|z| \leqq 1$. Then

$$U_n(\theta) + iV_n(\theta) = \sum_{k=0}^{n} d_{n,k} A_k e^{ik\theta},$$

and it follows that

$$f(z) - \sum_{k=0}^{n} d_{n,k} A_k z^k = O(n^{-\alpha})$$

uniformly for $|z| \leqq 1$. The extension to positive values of j is readily effected by methods similar to those employed by Jackson‡ in this connection; the details are left to the reader.

* A polynomial in z of degree n is any expression of the form $a_0 + a_1 z + a_2 z^2 + \cdots + a_n z^n$; we do not assume a_n to be different from zero.

† Jackson, *The Theory of Approximation*, Colloquium Publications of this Society, vol. 11, 1930, pp. 1–12; in particular, Theorem IV; Bernstein, Mémoires de l'Académie Royale de Belgique, (2), vol. 4 (1912), pp. 88–89.

‡ Jackson, op. cit., pp. 9–11.

We turn to the general case. Let $w=\psi_\nu(z)$ denote an analytic function which maps the interior of the curve C_ν conformally onto the interior of the circle γ: $|w|=1$, and let $z=\psi_\nu^{-1}(w)$ denote the inverse of this function. A number $R>1$ exists such that the analytic curves C_ν': $\psi_\nu(z)=R$, $(\nu=1, 2, \cdots, \lambda)$, are mutually exterior and such that the function $\psi_\nu(z)$ is regular and univalent in a limited simply connected region containing the curve C_ν', $(\nu=1, \cdots, \lambda)$. By using the fact that the function $\psi_\nu^{-1}(w)$ is regular for $|w|<R$, we may easily show that each of the functions $f_\nu[\psi_\nu^{-1}(w)]$ belongs to the class Lip (γ, j, α). Hence for each value of ν there exist functions $F_{\nu,n}(z)$ that are polynomials in $w=\psi_\nu(z)$ of respective degrees n, $(n=1, 2, \cdots)$, such that

$$f_\nu(z) - F_{\nu,n}(z) = O(n^{-j-\alpha}) \tag{1}$$

uniformly for z on and interior to C_ν. The functions $F_{\nu,n}(z)$, $(n=1, 2, \cdots)$, are regular in a limited simply connected region containing the curve C_ν'. Moreover, since $F_{\nu,n}(z)=O(1)$ uniformly for $|\psi_\nu(z)|=1$, it follows from a result of Faber* that $F_{\nu,n}(z)=O(R^n)$ uniformly for $|\psi(z)|\leqq R$. We define functions $F_n(z)$ as follows: $F_n(z)\equiv F_{\nu,n}(z)$, z on and interior to C_ν', $(n=1, 2, \cdots; \nu=1, 2, \cdots, \lambda)$. Then

$$F_n(z) = O(R^n) \tag{2}$$

uniformly for z on and interior to C_ν', $(\nu=1, \cdots, \lambda)$.

Now let $\phi=\phi(z)$ denote an analytic function which maps the complement of S conformally (but not necessarily uniformly) onto the region $|w|>1$ so that the point $z=\infty$ corresponds to the point $w=\infty$. We choose a number $\mu>1$ such that the locus Γ: $|\phi(z)|=\mu$ consists of λ analytic curves lying respectively interior to the curves C_ν' but containing the curves C_ν in their respective interiors.† We also select numbers μ_1, $1<\mu_1<\mu$, and r, $\mu_1/\mu<r<1$. The λ components of the locus Γ_1: $|\phi(z)|=\mu_1$ lie respectively interior to those of the locus Γ, but contain the curves C_ν in their respective interiors.

* Faber, Münchner Berichte, 1922, pp. 157–178. See also Walsh, *Interpolation and Approximation by Rational Functions in the Complex Domain*, Colloquium Publications of this Society, vol. 20, 1935, pp. 77–78.

† For the proof that such a number μ exists, see, for instance, Walsh, op. cit., pp. 65–68, where the locus $|\phi(z)|=K\geqq 1$ is described in some detail.

It has been established by Szegö* that there exists on the boundary of S a set of points $z_k^{(m)}$, $(k=1, 2, \cdots, m+1; m=0, 1, 2, \cdots)$, such that

$$|\omega_m(z)|^{1/m} \to \Delta\,|\phi(z)| \tag{3}$$

uniformly for z on any closed point set of the complement of S, where Δ denotes the transfinite diameter of S, and where $\omega_m(z)=(z-z_1^{(m)})(z-z_2^{(m)})\cdots(z-z_{m+1}^{(m)})$. The polynomial of degree m which coincides with the function $F_n(z)$ in the points $z_k^{(m)}$, $(k=1, 2, \cdots, m+1)$, may be written as follows:

$$P_m(z)=\frac{1}{2\pi i}\int_\Gamma \frac{F_n(t)}{t-z}\left(1-\frac{\omega_m(z)}{\omega_m(t)}\right)dt,$$

as may be verified directly; and for z on Γ_1 we have

$$|P_m(z)-F_n(z)| \leqq \frac{1}{2\pi}\int_\Gamma \left|\frac{F_n(t)}{t-z}\,\frac{\omega_m(z)}{\omega_m(t)}\right||dt|. \tag{4}$$

By (3) we have

$$\left|\frac{\omega_m(z)}{\omega_m(t)}\right| < r^m, \qquad (z \text{ on } \Gamma_1,\ t \text{ on } \Gamma), \tag{5}$$

for all values of m sufficiently large. Let d denote the minimum of $|t-z|$ for t on Γ, z on Γ_1, and let l denote the length of Γ. Let $m=qn$, where q is a positive integer such that $r_1=r^qR<1$. From (2), (4), and (5) we obtain

$$|P_{qn}(z)-F_n(z)| \leqq \frac{M_1R^n}{d}r^{qn}l = M_2r_1^n, \qquad (z \text{ on } \Gamma_1), \tag{6}$$

for all values of n sufficiently large, where M_1 and M_2 are independent of n and z. By the principle of the maximum, (6) holds for z interior to each of the curves which compose Γ_1. When we combine (6) with (1), we have

$$f(z)-P_{qn}(z)=O(n^{-i-\alpha}+r_1^n)=O(n^{-i-\alpha})$$

* Szegö, Mathematische Zeitschrift, vol. 21 (1924), pp. 205–207; see also Walsh and Russell, Transactions of this Society, vol. 36 (1934), pp. 13–28; Walsh, op. cit., pp. 68–75. The result was proved by Fejér (Göttinger Nachrichten, 1918, pp. 319–331) in the case $\lambda=1$, using methods developed by Hilbert (Göttinger Nachrichten, 1897, pp. 63–70).

uniformly for z on S. We now define the polynomials $p_n(z)$ referred to in the statement of the theorem by the following relations: $p_n(z) \equiv 0$, $(n = 1, 2, \cdots, q-1)$, $p_{qn'+h}(z) \equiv P_{qn'}(z)$, $(h = 0, 1, \cdots, q-1; n' = 1, 2, \cdots)$. Since $O(n'^{-j-\alpha}) = O[(qn'+h)^{-j-\alpha})]$, it follows that $f(z) - p_n(z) = O(n^{-j-\alpha})$ uniformly for z on S, and the proof is complete.

If the smoothness of the function $f(z)$ on the boundary of S is described in terms of a modulus of continuity instead of in terms of a Lipschitz condition, the above methods may be used to derive a result analogous to Theorem IV on page 12 of Jackson's *The Theory of Approximation*. The following special case is of importance.*

THEOREM. *If the function $f(z)$ is regular in each of the regions contained in S and is continuous and possesses a continuous jth derivative for z on S, then there exist polynomials $p_n(z)$ of respective degrees n, $(n = 1, 2, \cdots)$, such that $f(z) - p_n(z) = O(n^{-j})$ uniformly for z on S.*

The theorem may be used to obtain results on the degree of convergence of certain special sequences of functions. For example, let C be an analytic Jordan curve of the finite plane, let $f(z)$ be a function of the class Lip (C, j, α), and let polynomials $L_n(z)$ of respective degrees n, $(n = 0, 1, 2, \cdots)$, be defined by the requirement of coinciding with the function $f(z)$ in the points $z_k^{(n)} = \Phi(e^{2\pi i k/(n+1)})$, $(k = 1, 2, \cdots, n+1)$, where the function $z = \Phi(w)$ maps the unlimited region bounded by C onto the region $|w| > 1$ so that the point $w = \infty$ corresponds to the point $z = \infty$. A previous result of the author† expresses the degree of convergence of the sequence $\{L_n(z)\}$ in terms of that of an arbitrary sequence of polynomials. This result, when combined with the first theorem of the present paper, yields‡ the equations

---

* The theorem for the case $j = 0$ is included in a theorem of Walsh; see Walsh, op. cit., p. 47, Theorem 15.

† J. H. Curtiss, Transactions of this Society, vol. 38 (1935), pp. 458–473; p. 467.

‡ This application of the theorem is analogous to Jackson's application of the results in pp. 1–12 of his book to the determination of the degree of convergence of trigonometric sums interpolating to a real function $F(x)$ in equally spaced points. These sums exhibit the degree of convergence involved in (8)

$$f(z) - L_n(z) = O(n^{-j-\alpha}) \tag{7}$$

uniformly for z on an arbitrary closed point set T interior to C and

$$f(z) - L_n(z) = O(n^{-j-\alpha} \log n) \tag{8}$$

uniformly for z on C.

We may also obtain the degree of convergence of certain sequences of Riemann sums which tend to the Cauchy integral of the function $f(z)$ over the curve C. If we write

$$\sigma_n(z) = \sum_{k=1}^{n+1} \frac{f(z_k^{(n)})}{z - z_k^{(n)}} (z_{k+1}^{(n)} - z_k^{(n)}),$$

then it is possible to show that*

$$\sigma_n(z) - 2\pi i L_n(z) = O(\rho^n) \tag{9}$$

uniformly for z on T, where $\rho < 1$. Comparison of (7) with (9) indicates that†

$$\sigma_n(z) - \int_C \frac{f(t)}{t - z}\, dt = O(n^{-j-\alpha})$$

uniformly for z on T.

THE JOHNS HOPKINS UNIVERSITY

---

when the function $F(x)$ has properties corresponding to those of a function of class Lip (C, j, α).

W. E. Sewell (this Bulletin, vol. 41 (1935), pp. 111–117) has recently shown that the partial sums of the expansion of a function $f(z)$ which is of class Lip (C, j, α) in the Faber polynomials belonging to the region interior to C also exhibit the degree of convergence involved in (8) for z on C.

* This equation may be obtained by methods used by Curtiss, loc. cit.; see in particular pp. 460–464 and p. 466.

† For further results on degree of convergence of Riemann sums, see Pólya and Szegö, *Aufgaben und Lehrsätze*, vol. 1, 1925, pp. 35–37, 194–195; and see Walsh and Sewell, this Bulletin, vol. 42 (1936), p. 489, Abstract 284.

CONTINUITY IN TOPOLOGICAL GROUPS*

BY DEANE MONTGOMERY†

In the theory of topological groups it is customary to make certain assumptions concerning the continuity of the product and the continuity of the inverse. It will be shown here that for certain types of group spaces less stringent assumptions than those usually made yield the ordinary assumptions as theorems.

Suppose that G is a metric space‡ whose elements form a group. If x and y are any two elements of G, the distance between them will be denoted by $d(x, y)$ and their (group) product will be denoted by $x \cdot y$ or xy. The inverse of x will be denoted by x^{-1}, and the identity of the group by e. If H is a set of elements of G, then xH, Hx, and H^{-1} are sets in G having an obvious definition. The function xy is a function defined everywhere in the product space $G \times G$. It is often assumed that this function is continuous in the two variables simultaneously, but the following theorem shows that in a large class of cases the simultaneous continuity follows from continuity in each variable separately and this with no continuity restriction whatever on the inverse function. In fact it will be shown for separable groups that the continuity of the inverse also follows from the continuity of xy in x and y separately.

THEOREM 1. *If G is locally complete§ and the function xy is continuous in each variable separately, then it is continuous in the two variables simultaneously.*

Let it be noted first that it is sufficient to prove the simultaneous continuity at (e, e),‖ for if there is a discontinuity any-

* Presented to the Society, June 18, 1936.

† The results of this note were partially obtained when the author was a National Research Fellow at Princeton University and the Institute for Advanced Study. Some of the questions considered here were raised in a seminar conducted by W. Mayer.

‡ Fréchet introduced such spaces. For an account of them see his *Les Espaces Abstraits*. See also Kuratowski, *Topologie*, I.

§ G is said to be locally complete if there is about every point an open set whose closure is complete.

‖ The customary notation for points in a product space is followed here.

where there will be one here. In order to see this let a_n be a sequence approaching a and let b_n be a sequence approaching b, while a_nb_n does not approach ab. Because of the left and right continuity $a^{-1}a_n$ and b_nb^{-1} are sequences approaching e, but $a^{-1}a_nb_nb^{-1}$ does not approach e for if it did we could use first the left and then the right continuity to show that a_nb_n approaches ab.

It is convenient first to prove a lemma.

LEMMA. *If H is any open set in G and ϵ is any positive number, then there exists an open subset H_1 of H and a positive number δ such that for all elements h of H_1 and any element a of G the relation $d(a, e)<\delta$ implies the relation $d(ah, h)\leqq\epsilon$.*

Let B_n denote all elements h of G such that for any a, $d(a, e)<1/n$ implies $d(ah, h)\leqq\epsilon$. The set B_n is closed, a fact which may be seen as follows. Suppose that B_n is not closed and that b_m is a sequence of elements of B_n approaching an element b not in B_n. Since b is not in B_n, there is some element a such that $d(a, e)<1/n$ and (1) $d(ab, b)>\epsilon$. For all b_m, however, (2) $d(ab_m, b_m)\leqq\epsilon$. Because xy is continuous in y, $\lim ab_m=ab$. Thus (1) and (2) are contradictory and from this contradiction it may be concluded that B_n is closed.

Because of the left continuity of xy every h in H belongs to B_n for sufficiently large n. Therefore $H\subset\sum_nB_n$. Since H is of the second category,* there must be some n such that $H\cdot B_n$† is of the second category. Then B_n must be everywhere dense in some open subset H_1 of H; and from the fact that B_n is closed, B_n must include all of H_1. The lemma is now demonstrated.

The proof of Theorem 1 may now be given. Let G (for uniformity denote G by H_0) be the first open subset of G to which the lemma is applied; by this lemma there exists a positive number δ_1 and an open subset H_1 of H_0 such that for all elements a in G and all elements h in H_1, $d(a, e)<\delta_1$ implies that $d(ah,h)\leqq 1$.

Application of the lemma next to H_1 shows that there is a δ_2 and an open subset H_2 of H_1 such that for all elements a in G and all elements h in H_2, $d(a, e)<\delta_2$ implies $d(ah, h)\leqq 1/2$.

* Banach, *Théorie des Opérations Linéaires*, p. 14. By hypothesis H contains complete metric subspaces and therefore the statement follows from Banach's theorem at once.

† This denotes the intersection or point set product of H and B_n.

Proceeding in this manner, we obtain for every n a δ_n and an open subset H_n of H_{n-1} such that for all a in G and all h in H_n, $d(a, e) < \delta_n$ implies $d(ah, h) \leqq 1/n$. It may be assumed that the diameter of H_n is less than $1/n$, and that $\overline{H}_{n+1} \subset H_n$, and that for some n, H_n is complete, the last assumption being possible because G is locally complete. Under these conditions, we have $\prod_n \overline{H}_n = \prod_n H_n = h_0$, where h_0 is some point of G.

Let ϵ be any positive number whatever. Since xy is continuous in x and since h_0^{-1} is constant, there is a number δ such that, for any h in G, $d(h, h_0) < \delta$ implies $d(hh_0^{-1}, h_0h_0^{-1}) = d(hh_0^{-1}, e) < \epsilon$. Let n be so large that $1/(2n) < \delta/2$. By the definition of δ_i it is true for all h in H_{2n} that $d(a, e) < \delta_{2n}$ implies $d(ah, h) \leqq 1/(2n)$. Now let $S(e, \delta_{2n})$ be the open sphere of center e and radius δ_{2n} and let $O = [S(e, \delta_{2n})] \cdot [H_{2n} \cdot h_0^{-1}]$. The set O is open and includes e. If b is an element of O, $b < hh_0^{-1}$, where h is in H_{2n}. Let a be any other element of O. Then $d(ab, e) = d(ahh_0^{-1}, h_0h_0^{-1})$. But $d(ah, h) \leqq 1/(2n)$, and $d(h, h_0) < 1/(2n)$. Therefore $d(ah, h_0) < 1/n < \delta$, and it follows that

$$d(ab, e) = d(ahh_0^{-1}, h_0h_0^{-1}) = d(ahh_0^{-1}, e) < \epsilon.$$

Hence the function xy is continuous at (e, e), because for an arbitrary ϵ there has been found an open set $O \times O$ including $e \times e$ such that for any element (a, b) in $O \times O$, $d(ab, e) < \epsilon$. By the remark immediately following Theorem 1, it is evident that the proof is now complete.

This theorem could be easily proved if G were assumed to be separable, by making use of known theorems. Since xy is continuous in each variable separately, it is of Baire class 1 in the two variables together. It therefore has points of continuity* and if it has any points of continuity it is continuous everywhere, as can be seen from the remark immediately following the statement of the theorem. In the non-separable case xy is of class 1 as before but whether it has points of continuity does not follow in this case from any known theorem. It would be interesting to know whether or not the next theorem, which is proved for only the separable case, is also true in the non-separable case.

THEOREM 2. *If G is complete and separable and if xy is continuous in x and y separately, then x^{-1} is continuous.*

* See Kuratowski, loc. cit., pp. 180, 189, for the relevant theorems.

It follows from Theorem 1 that xy is continuous in x and y simultaneously. It will now be shown that x^{-1} is a function in the Baire classification. In order to do this it is sufficient to prove that if F is any closed set in G, then F^{-1} is a Borel set. Let M denote the set of points (x, y) of $G \times G$ such that $xy = e$. This set is closed because of the continuity of xy. Now let $N = (G \times F) \cap M$. The projection of this set on G is F^{-1}. This is because N contains those points of $G \times G$ which are of the form (x, y), where y is in F and $xy = e$. Further, no two points of N project into the same point so that F^{-1} is the continuous (1-1) image of N. Since N is not necessarily compact, it can not be concluded that F^{-1} is closed, but under the present circumstances it can be concluded that F^{-1} is a Borel set* and hence x^{-1} is a Baire function.

The proof of the theorem is now completed by a lemma.

LEMMA. *If G is separable and complete and if xy is continuous in each variable separately and x^{-1} is a Baire function, then x^{-1} is continuous.*

The proof of this lemma follows with little variation the proof of a theorem of Banach.† First note that it is sufficient to prove that x^{-1} is continuous in the neighborhood of e (see Banach). Since x^{-1} is a Baire function it is continuous on a set H, where $G - H$ is of the first category. Let a_n be a sequence of elements in G approaching e. Since $G - H$ is of the first category $a_n^{-1}(G - H)$ is also of the first category. It follows (see Banach) that $G - H \cdot \prod (a_n^{-1} H)$ can not equal G so that there is a point a in H and in $a_n^{-1} H$ for each n. Therefore $a_n a$ is in H. Since $a_n a$ approaches a and since x^{-1} is continuous on H, it follows that $(a_n a)^{-1} = a^{-1} a_n^{-1}$ approaches a^{-1} and that a_n^{-1} approaches e.

Theorem 2 clearly remains true if we replace the hypothesis of completeness by the hypothesis of local completeness.‡

SMITH COLLEGE

---

* Kuratowski, loc. cit., p. 251.

† Banach, loc. cit., p. 23.

‡ In fact, if a space is locally complete, it may be metrized so as to be complete; but for some applications the hypothesis given is more convenient.

LIE GROUPS SIMPLY ISOMORPHIC WITH NO LINEAR GROUP†

BY GARRETT BIRKHOFF

1. *Introduction.* One of the most interesting conjectures concerning finite continuous groups is the conjecture that every Lie group is topologically isomorphic with a group of matrices. The proof of this conjecture, even in the small, would establish the truth of the famous conjecture that every Lie group nucleus (or *germ*) is a piece of a Lie group.‡ This makes it of interest to know that there exist Lie groups in the large, simply isomorphic even in the purely algebraic sense—and a fortiori topologically isomorphic in the large—with no group of matrices. It is to the proof of this fact that the present note is devoted.

2. *The Basic Lemma.* The proof ultimately rests on the following lemma.

LEMMA 1. *Let Γ be any group of linear transformations. Suppose Γ contains elements S and T whose commutator $R=S^{-1}T^{-1}ST$ is of prime order p, and satisfies $SR=RS$, $TR=RT$. Then Γ is of degree at least p.*

By the *degree* of Γ is meant the *dimensions*—that is, the maximum number of linearly independent elements—of the linear space Σ on which Γ operates. The vector elements of Σ will be written $x, y, z, \cdots$.

PROOF. Since R is of order p, Σ contains a vector x such that $Rx=\alpha x$, where α is a primitive pth root of unity.§ Now let Λ denote the linear subspace of *all* vectors $x\epsilon\Sigma$ satisfying $Rx=\alpha x$. If $x\epsilon\Lambda$, then

$$R(Sx) = S(Rx) = S(\alpha x) = \alpha(Sx).$$

† Presented to the Society, September 3, 1936.

‡ For a fuller description of the status of these conjectures, see [3], especially §17. See also [1], p. 24, middle.

Numbers in square brackets refer to the Bibliography at the end of this paper.

§ This follows from [4], Theorem 125.

That is, $Sx \epsilon \Lambda$, showing that S, and likewise T, transforms Λ into itself.

But now observe that *within* Λ, $S^{-1}T^{-1}ST = \alpha I$ (where I denotes the identical matrix), whence $T^{-1}ST = S$. But $T^{-1}ST$ and S have (see [4], Theorem 126) the same characteristic roots; hence so do S and αS, as linear transformations of Λ. Moreover since S is non-singular, its characteristic roots are not zero; hence it has at least p characteristic roots

$$\mu_0, \alpha\mu_0, \cdots, \alpha^{p-1}\mu_0, \qquad (\mu_0 \neq 0).$$

It follows that Λ, and therefore Σ, has at least p dimensions.

3. *The Main Theorem.* With the help of this lemma, one can easily exhibit a Lie group† simply isomorphic with no linear group. Let G_3 denote the group of all matrices

$$M(x, y, z) = \begin{pmatrix} 1 & x & z \\ 0 & 1 & y \\ 0 & 0 & 1 \end{pmatrix}.$$

And let N denote the discrete normal subgroup of matrices $M(0, 0, n)$, n any integer. Finally, let $G_3^* = G_3/N$.

THEOREM 1. *G_3^* is a Lie group simply isomorphic with no linear group.*

PROOF. That G_3^* is a Lie group can be verified directly, and also follows from general principles (see [1], p. 12). It remains to show that it is simply isomorphic with no linear group. But let p be any prime; then the images S of $M(1, 0, 0)$ and T of $M(0, 1/p, 0)$ both commute with $R = S^{-1}T^{-1}ST$, which is the image of $M(0, 0, 1/p)$ in G_3, while R is of order p. Hence by Lemma 1 any linear group simply isomorphic with G_3^* would have to be of degree at least p, and since p can be *any* prime, this is absurd.

COROLLARY. *No group having G_3^* as a subgroup is simply isomorphic with a linear group.*

REMARK 1. Since G_3 was a linear group, we see that there can

† By a *Lie group* we shall mean a *connected finite continuous group* in the usual sense; *mixed* groups will be excluded. See [1], p. 7.

exist no construction yielding, from representations of a given group as a linear group, non-trivial representations of its homomorphic images as linear groups.

REMARK 2. As all the two-parameter Lie groups are known ([1], p. 25), and can be realized in the large as linear groups, G_3^* is the simplest example of a group isomorphic in the large with no linear group.

4. *Additional Definitions.* In §§4–6 we shall show how to find whole families of Lie groups simply isomorphic with no linear group. But in order to do this, we shall need to recall some familiar parts of group theory.

By the *central* of a Lie group G is meant the set of all elements of G commuting with all other elements; this is a topologically closed and invariant subgroup of G.

Every Lie group G has a *Lie algebra* A, defined from the set of the infinitesimal generators $U, V, \cdots$ of G by regarding commutation (the taking of Poisson brackets) as a non-associative multiplication. Two Lie groups are locally isomorphic if and only if they have isomorphic Lie algebras ([2], [32.2]).

We shall call G *hypercentral* if and only if the associated Lie algebra is *nilpotent*, in the sense that all products of sufficient length r reduce to zero. This terminology is justified since G is hypercentral if and only if, denoting by Z_1 the subgroup of G consisting of the identity 0, and by Z_{k+1}/Z_k inductively the central of G/Z_k, we have $Z_r = G$. (Actually, Z_k contains the invariant sub-algebra A_{r-k} of A consisting of linear combinations of products of lengths $\geqq (r-k)$.)

REMARK. Any Lie group of *triangular* matrices $A = \|a_{ij}\|$ satisfying $a_{ii} = 1$ and $a_{ij} = 0$ if $i > j$ is hypercentral. All Poisson brackets of length exceeding its degree vanish.

5. *A Theorem of Linear Groups.* We shall now prove a result of some intrinsic interest,† which will be used in the sequel.

THEOREM 2. *Let U, V, W be any three matrices which satisfy $UV - VU = W$, $UW = WU$, $VW = WV$. Then the matrices $M(\lambda, \lambda', \lambda'') = \exp(\lambda U + \lambda' V + \lambda'' W)$ form a group H, topologically isomorphic in the large with G_3.*

† Theorem 2 is especially interesting as U, V, W have the formal properties of the position, momentum, and scalar operators p, q, $qp - pq$ of quantum mechanics.

EXPLANATION. As is customary,† we use the notation

$$\exp X = I + X + \frac{1}{2!}X^2 + \frac{1}{3!}X^3 + \cdots .$$

PROOF. Since for any X, $[\exp X] \circ [\exp(-X)] = I$, and since in H one can prove the special identity

$$M(\lambda, \lambda', \lambda'') \circ M(\mu, \mu', \mu'') = M(\lambda + \mu, \lambda' + \mu', \lambda'' + \mu'' - \lambda'\mu), \tag{1}$$

clearly H is a group. Further, (1) exhibits a continuous homomorphism‡ of G_3 into H, since (1) expresses the rule for matrix multiplication in G_3.

We shall show that this is a topological isomorphism, in other words, that if $M(\lambda_n, \lambda_n', \lambda_n'') \equiv M_n \to M(\lambda, \lambda', \lambda'')$, then $\lambda_n \to \lambda$, $\lambda_n' \to \lambda'$, and $\lambda_n'' \to \lambda''$. By group-homogeneity, it is sufficient to prove this for the special case $\lambda = \lambda' = \lambda'' = 0$, $M(\lambda, \lambda', \lambda'') = I$. This is what we shall do.

The kernel of the demonstration is a well known theorem of Lie ([2], p. 163) which shows that one can so transform the matrices of H that the generating matrices $U = \|u_{ij}\|$ and $V = \|v_{ij}\|$ are *triangular*, that is, satisfy $u_{ij} = v_{ij} = 0$, if $i > j$. But now computing the matrix $W = UV - VU = \|w_{ij}\|$, we see that $w_{ij} = 0$ unless $(j-i) \geqq 1$, while it is easily verified to be a general property of matrix multiplication, that if Y and Z are any two matrices such that $y_{ij} = 0$ unless $(j-i) \geqq \alpha_1$, and $z_{ij} = 0$ unless $(j-i) \geqq \alpha_2$, and $X = YZ$, then $x_{ij} = 0$ unless $(j-i) \geqq \alpha_1 + \alpha_2$. It is a corollary that the non-zero coefficients of W nearest the principal diagonal, that is, minimizing $(j-i)$ subject to $w_{ij} \neq 0$, appear as $\lambda'' w_{ij}$ in $\exp(\lambda'' W)$, and so the subgroup of the $\exp(\lambda'' W)$ is topologically isomorphic in the large with the translation-group $x \to x + t$.

Unless $\lambda_n \to 0$, we can, taking a suitable subsequence, assume without loss of generality that $1/\lambda_n = \bar{\lambda}_n \to \bar{\lambda}$. But irrespective of n,

$$M_n^{-1} \circ \exp(-\bar{\lambda}_n V) \circ M_n \circ \exp(\bar{\lambda}_n V) = \exp W \neq I;$$

† See J. von Neumann, Mathematische Zeitschrift, vol. 30 (1929), pp. 3–42.

‡ For the distinction between continuous homomorphisms and topological isomorphisms see H. Freudenthal, *Einige Sätze über topologische Gruppen*, Annals of Mathematics, vol. 37 (1936), p. 46.

while on the other hand, $M_n \to I$, $\exp(\bar{\lambda}_n V) \to \exp(\bar{\lambda} V)$, and $I^{-1} \circ \exp(-\bar{\lambda} V) \circ I \circ \exp(\bar{\lambda} V) = I$. Since group multiplication and passage to the inverse are continuous operations, this is absurd.

This proves $\lambda_n \to 0$; similarly $\lambda_n' \to 0$. Hence it follows that $\exp(\lambda_n'' W) \to I$, whence by the preceding paragraph $\lambda_n'' \to 0$, which completes the proof.

REMARK. It is a corollary of §5 alone that G_3^* is topologically isomorphic in the large with no linear group. (For its generators satisfy the hypotheses of Theorem 2.)

6. *The Extended Theorem.* Using Theorems 1 and 2, one can exhibit without difficulty a whole class of Lie groups simply isomorphic with no linear group. First we note an almost obvious lemma.

LEMMA 2. *Any non-Abelian hypercentral Lie group G has infinitesimal generators U, V, W satisfying $[U, V] = W$ and $[X, W] = 0$ for every X.*

PROOF. Since G is not Abelian, U_0 and V_0 exist satisfying $[U_0, V_0] = V_1 \neq 0$. Moreover either $[X, V_1] = 0$ for every X, or $[U_1, V_1] = V_2 \neq 0$ for some infinitesimal generator U_1. But since the Lie algebra of G is nilpotent, the latter alternative cannot recur indefinitely, and so we finally get elements U_q and V_q such that $[U_q, V_q] = V_{q+1}$, while $[X, V_{q+1}] = 0$ for every X.

LEMMA 3. *Any non-Abelian hypercentral Lie group G of matrices has a subgroup S_3 whose central lies in the central of G, topologically isomorphic in the large with G_3.*

PROOF. By Lemma 1, the U, V, and W of Lemma 2 generate a subgroup S_3 topologically isomorphic in the large with G_3, whose central therefore consists of the $\exp(\lambda W)$. But since the $\exp(\lambda W)$ commute with all the infinitesimal generators X of G, and G (being connected) is generated by these, the $\exp(\lambda W)$ lie in the central of G.

THEOREM 3. *Any non-Abelian hypercentral Lie group G of matrices is locally topologically isomorphic with a Lie group G/N which is simply isomorphic with no linear group.*

PROOF. Refer to Lemma 3, and let N be the subgroup of the

exp (nW), n any integer. Then N is a discrete subgroup of the central of G, and so (see [1], p. 12) G/N is a Lie group locally topologically isomorphic with G.

But the homomorphism $G \to G/N$ carries S_3 into S_3/N, which is simply isomorphic with $G_3/N = G_3^*$. This and the corollary to Theorem 1 complete the proof.

E. Cartan [5] has shown that the universal covering group of the group of projective transformations of the line is topologically isomorphic in the large with no linear group.

BIBLIOGRAPHY

[1] E. Cartan, *Théorie des groupes finis et continus et l'analysis situs*, Mémorial des Sciences Mathématiques, no. 42, 1930.

[2] L. P. Eisenhart, *Continuous Groups of Transformations*, 1933.

[3] W. Mayer and T. Y. Thomas, *Foundations of the theory of continuous groups*, Annals of Mathematics, vol. 36 (1935), pp. 770–822.

[4] A. Speiser, *Gruppentheorie*, 2d ed., 1927.

[5] *La Topologie des Groupes de Lie*, Paris, 1936, p. 18.

SOCIETY OF FELLOWS, HARVARD UNIVERSITY

CHARACTERISTICS OF BIRATIONAL TRANSFORMS IN S_r

BY B. C. WONG

1. *Introduction.* Consider a k-dimensional variety, V_k^n, of order n in an r-space, S_r. Let us project V_k^n from a general $(r-k-t-1)$-space of S_r upon a general $(k+t)$-space of S_r and denote the projection by ${}_tV_k^n$. We are supposing that $1 \leqq t \leqq k$. Then upon ${}_tV_k^n$ lies a double variety, D_{k-t}, of dimension $k-t$ and order b_t and upon D_{k-t} lies a pinch variety, W_{k-t-1}, of dimension $k-t-1$ and order j_{t+1}. Since the symbol W_{-1} is without meaning, we thus obtain $2k-1$ characteristics $b_1, b_2, \cdots, b_k, j_2, j_3, \cdots, j_k$. The symbol j_1 has a meaning which will be explained subsequently.

Now let a general $(r-k+q-2)$-space, $S_{r-k+q-2}$, $(1 \leqq q \leqq k)$, be given in S_r. Through this $S_{r-k+q-2}$ pass ∞^{k-q+1} primes of S_r and ∞^{k-q} of these are tangent to V_k^n. The points of contact form a $(k-q)$-dimensional variety, U_{k-q}. Denote its order by m_q. Thus

we obtain k further characteristics $m_1, m_2, \cdots, m_k$. If we project V_k^n upon a $(k+1)$-space, S_{k+1}, of S_r, we see that m_q is the class of the V_q^n in which a $(q+1)$-space of S_{k+1} meets the projected variety. We also say that m_q is the class of the q-dimensional variety in which an $(r-k+q)$-space of S_r meets V_k^n.

In the case where V_k^n is the complete intersection of $r-k$ general primals, of orders $n_1, n_2, \cdots, n_{r-k}$, respectively, in S_r, the values of j_t, b_t, m_q are known* and they are

$$\text{(I)} \qquad j_t = n_1 n_2 \cdots n_{r-k} \sum (n_1 - 1)(n_2 - 1) \cdots (n_t - 1),$$

$$\begin{aligned}\text{(II)} \qquad b_t &= \frac{1}{2} n_1 n_2 \cdots n_{r-k} [n_1 n_2 \cdots n_{r-k} - 1 - \sum (n_1 - 1) \\ &\quad - \sum (n_1 - 1)(n_2 - 1) - \cdots \\ &\quad - \sum (n_1 - 1)(n_2 - 1) \cdots (n_t - 1)] \\ &= \frac{1}{2} n_1 n_2 \cdots n_{r-k} [\sum (n_1 - 1)(n_2 - 1) \cdots (n_{t+1} - 1) \\ &\quad + \sum (n_1 - 1)(n_2 - 1) \cdots (n_{t+2} - 1) + \cdots \\ &\quad + \sum (n_1 - 1)(n_2 - 1) \cdots (n_{r-k} - 1)],\end{aligned}$$

$$\text{(III)} \qquad m_q = n_1 n_2 \cdots n_{r-k} \sum_h \sum (n_1 - 1)^{h_1} (n_2 - 1)^{h_2} \cdots (n_q - 1)^{h_q},$$

where

$$h = h_1 + h_2 + \cdots + h_q = q.$$

We shall refer to these values later.

In this paper we propose to determine the values of the same characteristics for the variety V_k^n in S_r which we consider as the birational transform of a k-dimensional variety, say Φ_k^ν, of order ν in a ρ-space, Σ_ρ, for $\rho < r$. We confine ourselves to the case where Φ_k^ν is the complete intersection of $\rho - k$ general primals of Σ_ρ, of respective orders $\nu_1, \nu_2, \cdots, \nu_{\rho-k}$, given by the equations

$$\text{(1)} \qquad F^{(1)} = 0, \; F^{(2)} = 0, \cdots, F^{(\rho-k)} = 0,$$

* C. Segre, *Mehrdimensionale Räume*, Encyklopädie der Mathematischen Wissenschaften, III C7, pp. 944–945. Also B. C. Wong, *On certain characteristics of k-dimensional varieties in r-space*, this Bulletin, vol. 38 (1932), pp. 725–730.

where $F^{(i)}$ is a homogeneous function of degree ν_i in the variables $\xi_0, \xi_1, \cdots, \xi_\rho$. The order of Φ_k^ρ is $\nu = \nu_1\nu_2 \cdots \nu_{\rho-k}$. The corresponding characteristics η_i, β_i, μ_q of Φ_k^ρ are given by (I), (II), (III) if we replace in the right-hand members r by ρ and n_i by ν_i.

We suppose that the transformation of Φ_k^ρ into V_k^n is accomplished by means of a general linear ∞^r-system, $|\psi|$, without base varieties of any kind, of $(k-1)$-dimensional varieties of order νN, and that $|\psi|$ is the intersection of Φ_k^ρ and a general linear ∞^r-system, $|\phi|$, of primals of order N, none passing through Φ_k^ρ, given by the equation

$$a_0\phi^{(0)} + a_1\phi^{(1)} + \cdots + a_r\phi^{(r)} = 0, \tag{2}$$

the ϕ's being linearly independent homogeneous polynomials of degree N in the $(\rho+1)$ ξ's. Then, the order of V_k^n is $n = \nu N^k = \nu_1\nu_2 \cdots \nu_{\rho-k}N^k$. The coordinates of the points on V_k^n are given by

$$\sigma x_0 = \phi^{(0)}(\xi_0, \xi_1, \cdots, \xi_\rho), \qquad \sigma x_1 = \phi^{(1)}(\xi_0, \xi_1, \cdots, \xi_\rho),$$
$$\cdots, \qquad \sigma x_r = \phi^{(r)}(\xi_0, \xi_1, \cdots, \xi_\rho),$$

where the ξ's satisfy equations (1).

It is to be noted that an h-dimensional locus of order l on Φ_k^ρ goes into an h-dimensional locus of order lN^h on V_k^n. For $h = k$, Φ_k^ρ goes into V_k^n, where $n = \nu N^k$.

We shall first, in §2, derive a general relation connecting the b's and the j's for a general variety which has no extraordinary singular points. The determination of the values of the m's of our variety V_k^n will be given in §3 and the determination of those of the j's in §4. The values of the b's will then be obtained with the aid of the relation derived in §2. Incidentally, we find it interesting to express the m's and j's in terms of the μ's and η's, respectively, of Φ_k^ρ.

2. *The Relation between the b's and the j's.* Let C^l be a curve of order l, in a space of dimension greater than 2, whose points are paired in an (irrational) involution I_2. Suppose that C^l has d actual nodes at each of which two corresponding points of I_2 coincide but lie on different branches of the curve. If i denotes the number of simple points of C^l at each of which two corresponding points of I_2 become united, the order of the ruled sur-

face, which may be a cone, whose generators are lines joining corresponding points of the involution is, as is well known,

$$R=(2l-i-2d)/2.$$

Now consider a general k-dimensional variety V_k, of any order n, without extraordinary singular points, in S_r and let it be intersected by a general $(r-k+t)$-space of S_r in a V_t. If we project V_t upon a $(2t-1)$-space, S_{2t-1}, we see that the projection ${}_{t-1}V_t$ has a double curve D_1 of order b_{t-1} and j_t pinch points. This ${}_{t-1}V_t$ may certainly be regarded as the projection of a ${}_tV_t$ in a $(2t)$-space, S_{2t}, the ${}_tV_t$ being assumed to be a projection of V_t. Let Z be the point, taken in a general position of S_{2t}, from which ${}_tV_t$ is projected into ${}_{t-1}V_t$ in S_{2t-1}. There are ∞^1 lines through Z meeting ${}_tV_t$ in two distinct points and the locus of these lines is a ruled surface, in fact a cone, of order b_{t-1}. This cone meets ${}_tV_t$ in a curve c of order $2b_{t-1}$, of which the double curve D_1 on ${}_{t-1}V_t$ is the projection. The curve c has b_t actual nodes which are the improper double points of ${}_tV_t$. There are j_t elements of the cone tangent to ${}_tV_t$ and also to c. The projections of the points of contact are the pinch points on ${}_{t-1}V_t$. Now on c is an involution of pairs of points set up by the elements of the cone. There are b_t points each of which is the union of two corresponding points on different branches of the curve and j_t points each of which is the union of two corresponding points on a simple branch of the curve. Putting $R=b_{t-1}$, $l=2b_{t-1}$, $i=j_t$, $d=b_t$ in the relation of the paragraph just preceding, we have the desired relation $b_{t-1}=(4b_{t-1}-j_t-2b_t)/2$, or

$$2b_{t-2}=j_t+2b_t. \tag{3}$$

By letting $t=1, 2, \cdots, k$, successively, we obtain

$$2b_0=j_1+2b_1,\quad 2b_1=j_2+2b_2,\quad \cdots,\quad 2b_{k-1}=j_k+2b_k.$$

As we shall see presently, $b_0=n(n-1)/2$ and j_1 is identical with the class m_1 of a plane section of the projection ${}_1V_k$ in a $(k+1)$-space. The relation (3) may be replaced by the relation

$$2b_s-2b_t=j_{s+1}+j_{s+2}+\cdots+j_t, \qquad (s>t). \tag{4}$$

Note that this relation, or relation (3), is satisfied by (I) and (II).

3. *The Determination of the m's.* Returning to the V_k^n which is the birational transform of Φ_k^r in Σ_ρ, we see at once that m_q, $(1 \leqq q \leqq k)$, is the order of the variety U_{k-q} on V_k^n which has for image on Φ_k^r the complete intersection Θ_{k-q} of Φ_k^r and the Jacobian variety $\omega_{\rho-q}$ of the $\rho-k$ primals, given by (1), intersecting in Φ_k^r, and any $k-q+2$ independent primals of the system $|\phi|$, say $\phi^{(1)}=0, \phi^{(2)}=0, \cdots, \phi^{(k-q+2)}=0$. Θ_{k-q} is the locus of the points of contact between Φ_k^r and the ∞^{k-q} primals of the ∞^{k-q+1}-system determined by the $k-q+2$ primals just mentioned which are tangent to Φ_k^r. The conditions of contact are given by

$$\begin{Vmatrix} F_0^{(1)} & F_1^{(1)} & \cdots & F_\rho^{(1)} \\ \cdot & \cdot & \cdots & \cdot \\ F_0^{(\rho-k)} & F_1^{(\rho-k)} & \cdots & F_\rho^{(\rho-k)} \\ \phi_0^{(1)} & \phi_1^{(1)} & \cdots & \phi_\rho^{(1)} \\ \cdot & \cdot & \cdots & \cdot \\ \phi_0^{(k-q+2)} & \phi_1^{(k-q+2)} & \cdots & \phi_\rho^{(k-q+2)} \end{Vmatrix} = 0,$$

where $F_i^{(h)}$, $\phi_i^{(h)}$ are written in place of $\partial F^{(h)}/\partial \xi_i$, $\partial \phi^{(h)}/\partial \xi_i$, respectively. This equality represents the Jacobian $\omega_{\rho-q}$. The matrix being of $\rho+1$ columns and $\rho-q+2$ rows, the order of $\omega_{\rho-q}$ is*

$$H_q = \sum_{i=0}^{q} \binom{k+1-i}{q-i} (N-1)^{q-i} \cdot \sum_i \sum (\nu_1 - 1)^{h_1} (\nu_2 - 1)^{h_2} \cdots (\nu_i - 1)^{h_i}, \tag{5}$$

where

$$h_1 + h_2 + \cdots + h_i = i.$$

Then the order of Θ_{k-q} is $\nu_1\nu_2 \cdots \nu_{\rho-k} H_q$ and therefore the order of U_{k-q} on V_k^n which is the transform of Θ_{k-q} is

$$m_q = \nu_1\nu_2 \cdots \nu_{\rho-k} N^{k-q} H_q. \tag{6}$$

It is of interest to express m_q in terms of the μ's belonging to Φ_k^r. The various values of the μ's are obtained from (III) by changing r to ρ and n_i to ν_i. By taking account of (5), we have, from (6), writing μ_0 in place of ν or $\nu_1\nu_2 \cdots \nu_{\rho-k}$,

* Salmon, *Modern Higher Algebra*, 4th ed., Lesson 19.

$$m_q = N^{k-q}\sum_{i=0}^{q}\binom{k+1-i}{q-i}(N-1)^{q-i}\mu_i$$

$$= N^{k-q}\left[\binom{k+1}{q}(N-1)^q\mu_0 + \binom{k}{q-1}(N-1)^{q-1}\mu_1 + \cdots + \mu_q\right].$$

4. *The Determination of the j's and the b's.* The projection ${}_{t-1}V_k^n$, in a $(k+t-1)$-space of S_r, of V_k^n has a pinch variety W_{k-t} of order j_t. This W_{k-t} has for image on Φ_k^r the complete intersection of Φ_k^r and the variety $\pi_{\rho-t}$ common to all the primals each of which is the Jacobian of $\rho+1$ of the following $\rho+t$ primals:

$$F^{(1)}=0, \quad F^{(2)}=0, \quad \cdots, \quad F^{(\rho-k)}=0,$$

$$\phi^{(1)}=0, \quad \phi^{(2)}=0, \quad \cdots, \quad \phi^{(k+t)}=0.$$

The $k+t$ primals represented by the ϕ's equated to zero are supposed to be independent members of the system $|\phi|$ given by (2). The equations of $\pi_{\rho-t}$ are

$$\begin{Vmatrix} F_0^{(1)} & F_1^{(1)} & \cdots & F_\rho^{(1)} \\ \cdot & \cdot & \cdot & \cdot \\ F_0^{(\rho-k)} & F_1^{(\rho-k)} & \cdots & F_\rho^{(\rho-k)} \\ \phi_0^{(1)} & \phi_1^{(1)} & \cdots & \phi_\rho^{(1)} \\ \cdot & \cdot & \cdot & \cdot \\ \phi_0^{(k+t)} & \phi_1^{(k+t)} & \cdots & \phi_\rho^{(k+t)} \end{Vmatrix} = 0.$$

The left-hand member being a matrix of $\rho+t$ rows and $\rho+1$ columns, the order of $\pi_{\rho-t}$ is*

$$C_t = \sum_{i=0}^{t}\binom{k+t}{i}(N-1)^i\sum(\nu_1-1)(\nu_2-1)\cdots(\nu_{t-i}-1)$$

$$= \sum_{i=0}^{t}\binom{k+t}{t-i}(N-1)^{t-i}\sum(\nu_1-1)(\nu_2-1)\cdots(\nu_i-1), \tag{7}$$

* Salmon, loc. cit.

and the order of the intersection of Φ_k^r and $\pi_{\rho-t}$ is νC_t. Therefore, the order of W_{k-t} is

$$
\begin{aligned}
j_t &= \nu N^{k-t} C_t \\
&= \nu_1 \nu_2 \cdots \nu_{\rho-k} N^{k-t} \sum_{i=0}^{t} \binom{k+t}{t-i} (N-1)^{t-i} \\
&\qquad \cdot \sum (\nu_1 - 1)(\nu_2 - 1) \cdots (\nu_i - 1).
\end{aligned} \tag{8}
$$

Thus, for $t=1$, we have

$$
j_1 = \nu_1 \nu_2 \cdots \nu_{\rho-k} N^{k-1} [(k+1)(N-1) + \sum (\nu_1 - 1)] = m_1.
$$

Now in terms of the η's of Φ_k^r we have, from (I) and (8),

$$
\begin{aligned}
j_t &= N^{k-t} \sum_{i=0}^{t} \binom{k+t}{t-i} (N-1)^{t-i} \eta_i \\
&= N^{k-t} \left[\binom{k+t}{t} (N-1)^t \eta_0 + \binom{k+t}{t-1} (N-1)^{t-1} \eta_1 + \cdots \right. \\
&\qquad \left. + \binom{k+t}{1} (N-1) \eta_{t-1} + \eta_t \right],
\end{aligned}
$$

where $\eta_0 = \nu = \nu_1 \nu_2 \cdots \nu_{\rho-k}$.

To determine the b's we make use of the relation (3) or (4) of §2. A little calculation yields

$$
b_t = \frac{1}{2} \nu \left[\nu N^{2k} - \sum_{h=0}^{t} N^{k-h} C_h \right],
$$

where C_h is given by (7) if we replace in it t by h. Since any $(k-t)$-dimensional variety of order l on Φ_k^r goes into a $(k-t)$-dimensional variety of order lN^{k-t} on V_k^n, we see that

$$
\frac{2b_t}{N^{k-t}} = \nu^2 N^{k+t} - \nu \sum_{h=0}^{t} N^{t-h} C_h
$$

is the order of the image on Φ_k^r of the double variety D_{k-t} on the projection ${}_tV_k^n$ in a $(k+t)$-space.

THE UNIVERSITY OF CALIFORNIA

A MIXTURE THEOREM FOR NONCONSERVATIVE MECHANICAL SYSTEMS*

BY A. H. COPELAND†

We shall be concerned with a mechanical system which is acted upon by friction in such a manner that it eventually comes to rest. We shall assume that there is a probability distribution for the initial conditions of the system. There is then determined a corresponding distribution for the rest position. We shall show that in general as the friction approaches zero, this latter distribution approaches uniformity irrespective of the initial distribution. This type of problem was first solved by Poincaré.‡ The theory was later developed by Smoluchowski§ and recently an extensive contribution based on the ergodic theorem‖ has been made by Hopf.¶

We shall let $x_1, x_2, \cdots, x_\nu$ denote the position coordinates of the system at any time t and shall let x denote the vector $x_1, x_2, \cdots, x_\nu$. Thus dx/dt is the velocity vector. The initial position will be denoted by the vector B and the initial velocity by the vector V. Then V, B can be any point of 2ν-dimensional space Ω. We assume that there exists a continuous function ψ such that the probability of the initial conditions being represented by a point V, B in a given region S is equal to the integral of ψ taken over S. This condition is expressed by the equation

$$\text{prob } [(V, B)\epsilon S] = \int_S \psi(V, B)d\omega. \tag{1}$$

* Presented to the Society, December 27, 1934.

† Guggenheim fellow.

‡ *Calcul des Probabilités*, 1912.

§ *Über den Begriff des Zufalls und den Ursprung der Wahrscheinlichkeitsgesetze in der Physik*, Die Naturwissenschaften, vol. 6 (1918).

‖ Birkhoff, *Proof of a recurrence theorem for strongly transitive systems, Proof of the ergodic theorem*, Proceedings of the National Academy of Sciences, vol. 17 (1931).

¶ *On causality, statistics, and probability*, Journal of Mathematics and Physics, Massachusetts Institute of Technology, vol. 13 (1934). See also Struik, *On the foundations of the theory of probability*, Philosophy of Science, vol. 1 (1934).

Since ψ is a distribution function, it must satisfy the relation

$$\int_{\Omega} \psi(V, B)d\omega = 1. \tag{2}$$

Otherwise ψ will be unrestricted.

We shall assume that the coordinates $a_1, a_2, \cdots, a_\nu$ of the rest position are angular variables. That is, we shall be interested in the values of these variables reduced by some modulus. By a proper choice of units, we can make this modulus equal to 1. Thus we shall study the probability distribution for the vector $a-[a]$ whose components are $a_1-[a_1], a_2-[a_2], \cdots, a_\nu-[a_\nu]$, where $[a_i]$ is the largest integer contained in a_i. The vector $a-[a]$ lies in the region Δ given by the inequalities $0 \leqq z_i < 1$, where $i=1, 2, \cdots, \nu$. The probability distribution of $a-[a]$ will be called uniform if for every region E of Δ, we have the equation

$$\text{prob}(a-[a] \epsilon E) = m(E), \tag{3}$$

where $m(E)$ is the ν-dimensional Lebesgue measure of E. The region E may be any set whose frontier points constitute a set of measure 0. Our problem is to show that this uniform distribution is approached as the friction is decreased.

We shall first consider a system defined by the equation

$$\frac{d^2x}{dt^2} = \mu \cdot f\left(\frac{dx}{dt}\right), \tag{4}$$

where μ is a scalar parameter. If we substitute $x=B+\xi/\mu$ and $t=\tau/\mu$, then we observe that $dx/dt=d\xi/d\tau$ and $d^2x/dt^2=\mu d^2\xi/d\tau^2$. Hence

$$\frac{d^2\xi}{d\tau^2} = f\left(\frac{d\xi}{d\tau}\right). \tag{5}$$

We can obtain a solution $\xi=X(V, \tau)$, $d\xi/d\tau=\dot{X}(V, \tau)$ such that $X(V, 0)=0$ and $\dot{X}(V, 0)=V$. After a time T, the system (5) comes to rest. Thus $X(V, T)=A$ and $\dot{X}(V, T)=0$, where A is the rest position vector. If we solve one of the equations $\dot{X}(V, T)=0$ for T in terms of V and substitute this value in the equations $X(V, T)=A$, we obtain the relations $A=F(V)$

and $V=F^{-1}(A)$. We shall assume that system (4) (or (5)) is such that the Jacobian $J\binom{V}{A}$ is continuous and non-vanishing. Then

$$\int_S \psi(V, B)d\omega = \int_R \phi(A, B)d\sigma = \text{prob } [(A, B)\epsilon R], \tag{6}$$

where $\phi(A, B)=\psi[F^{-1}(A), B]\cdot J\binom{V}{A}$ and R is the transformed region S.

It will be observed that $x=B+X(V, \mu t)/\mu$, $dx/dt=\dot{X}(V, \mu t)$ is the solution of system (4) whose initial conditions are represented by the point V, B of the phase space Ω. The rest position of this system is the vector $a=B+A/\mu$. In order to study the distribution of $a-[a]$, we shall introduce a vector $b=-B/\mu+A$ and two systems of regions

$$\begin{aligned} r_{m,n}&: \quad m+\alpha \leqq a < m+\beta, \qquad n \leqq b < n+1, \\ R_{m,n}&: \quad m \leqq a < m+1, \qquad n \leqq b < n+1, \end{aligned} \tag{7}$$

where m and n stand respectively for the sets of integers $m_1, m_2, \cdots, m_\nu$ and $n_1, n_2, \cdots, n_\nu$. Then the set of all points A, B for which $\alpha \leqq a-[a]<\beta$ is $\sum_{m,n} r_{m,n}$, where the summation is extended over integral values of $m_1, m_2, \cdots, m_\nu, n_1, n_2, \cdots, n_\nu$. Hence

$$\begin{aligned} \text{prob } (\alpha \leqq a-[a] < \beta) &= \sum_{m,n} \text{prob } [(A, B)\epsilon r_{m,n}] \\ &= \sum_{m,n} \int_{r_{m,n}} \phi(A, B)d\sigma. \end{aligned} \tag{8}$$

The set of all points A, B is the region $\Gamma=\sum_{m,n} R_{m,n}$.

We shall apply the mean value theorem to evaluate the integrals occurring in this latter summation. To accomplish this, it will be necessary to compute the volumes of the regions $r_{m,n}$. An equation $b_i=-B_i/\mu+A_i$, where b_i is a constant, represents a line in the A_iB_i plane. If θ is the angle between such a line and the positive A_i axis, then $\tan\theta=\mu$. In a similar manner, an equation $a_i=B_i+A_i/\mu$, where a_i is a constant, is a line with slope $-1/\mu$. Hence the two inequalities $m_i+\alpha_i \leqq a_i < m_i+\beta_i$, $n_i \leqq b_i < n_i+1$, determine a rectangle whose area is

$$(\beta_i-\alpha_i)\cdot\sin^2\theta = (\beta_i-\alpha_i)\left(\frac{\mu^2}{1+\mu^2}\right).$$

It follows that the volume of $r_{m,n}$ is

$$\prod_{i=1}^{\nu} (\beta_i - \alpha_i) \cdot \left(\frac{\mu^2}{1 + \mu^2}\right)^{\nu}.$$

In a similar manner it can be seen that the volume of $R_{m,n}$ is $[\mu^2/(1+\mu^2)]^{\nu}$. It follows from the mean value theorem that

$$\begin{aligned} &\text{prob } (\alpha \leqq a - [a] < \beta) \\ &\quad = \sum_{m,n} \phi(\overline{A}_{m,n}, \overline{B}_{m,n}) \prod_{i=1}^{\nu} (\beta_i - \alpha_i) \left(\frac{\mu^2}{1 + \mu^2}\right)^{\nu} \\ &\quad = \prod_{i=1}^{\nu} (\beta_i - \alpha_i) \sum_{m,n} \phi(\overline{A}_{m,n}, \overline{B}_{m,n}) \left(\frac{\mu^2}{1 + \mu^2}\right)^{\nu}, \end{aligned} \tag{9}$$

where $\overline{A}_{m,n}$, $\overline{B}_{m,n}$ is a properly chosen point of the region $r_{m,n}$. Since $\overline{A}_{m,n}$, $\overline{B}_{m,n}$ is also a point of $R_{m,n}$ and since $[\mu^2/(1+\mu^2)]^{\nu}$ is the volume of this region, it follows from the definition of the definite integral that

$$\begin{aligned} \lim_{\mu \to 0} \text{prob } (\alpha \leqq a - [a] < \beta) &= \prod_{i=1}^{\nu} (\beta_i - \alpha_i) \int_{\Gamma} \phi(A, B) d\sigma \\ &= \prod_{i=1}^{\nu} (\beta_i - \alpha_i). \end{aligned} \tag{10}$$

Thus we have proved that

$$\lim_{\mu \to 0} \text{prob } (a - [a] \epsilon E) = m(E), \tag{11}$$

if E is the rectangular domain $\alpha \leqq z < \beta$. It remains to show that equation (11) holds for any set E whose frontier points are of measure 0. Corresponding to any such set and to any given positive number ϵ, there exist two sets E_1 and E_2 such that each consists of a finite sum of rectangular regions, that $E_1 < E < E_2$, and that $m(E_1 - E_2) < \epsilon$.* Hence it is easily seen that equation (11) holds for such sets.

Next we shall consider any mechanical system defined by second order differential equations involving a parameter μ.

* For a proof of this statement, see the author's article *Admissible numbers in the theory of geometrical probability*, American Journal of Mathematics, vol. 53 (1931).

The solution of these equations has the form $x=X(V, B, t, \mu)$, $dx/dt=\dot{X}(V, B, t, \mu)$, where $X(V, B, 0, \mu)=B$ and $\dot{X}(V, B, 0, \mu)=V$. We shall suppose that after a time T the system comes to rest. Thus $X(V, B, T, \mu)=a$ and $\dot{X}(V, B, T, \mu)=0$. Solving one of the equations $\dot{X}(V, B, T, \mu)=0$ for T and substituting this value in the function $X(V, B, t, \mu)$, we obtain the relations $a=F(V, B, \mu)$ and $V=F^{-1}(a, B, \mu)$. We shall suppose that the differential system is such that the limit as μ approaches 0 of $\mu \cdot F(V, B, \mu)$ exists and is not equal to 0. It is then natural to make the substitution $u=a\mu=\mu F(V, B, \mu)=f(V, B, \mu)$. We shall let $A=F(V, B, 1)$ and hence $V=F^{-1}(A, B, 1)$. Substituting this value of V in the function $F(V, B, \mu)$, we obtain the relations $a=G(A, B, \mu)$ and $u=g(A, B, \mu)$. We shall define a second system of equations $v=h(A, B, \mu)$ and shall assume that the functions h can be so determined that the partial derivatives $\partial A_i/\partial u_j$, $\partial A_i/\partial v_j$, $\partial B_i/\partial u_j$, $\partial B_i/\partial v_j$ are bounded and uniformly continuous throughout the range of the points u, v and for all values of μ in the interval $0 \leqq \mu \leqq 1$. Then $J\binom{A, B}{u, v}$ is bounded and uniformly continuous. We shall assume also that this Jacobian is non-vanishing.

We shall again define two systems of regions

$$
\begin{aligned}
(12)\qquad r_{m,n}&: \quad \mu(m+\alpha) \leqq u < \mu(m+\beta), \qquad \mu n \leqq v < \mu(n+1),\\
R_{m,n}&: \quad \mu m \leqq u < \mu(m+1), \qquad \mu n \leqq v < \mu(n+1).
\end{aligned}
$$

The diameters of $R_{m,n}$ and $r_{m,n}$ (considered as regions of the space Γ) approach 0 with μ. We have the equation

$$
\begin{aligned}
&\text{prob } (\alpha \leqq a - [a] < \beta) = \sum_{m,n} \int_{r_{m,n}} \phi(A, B)\, d\sigma \\
(13)\qquad &= \sum_{m,n} \phi(\overline{A}_{m,n}, \overline{B}_{m,n}) \cdot J\begin{pmatrix} A, & B \\ \bar{u}_{m,n}, & \bar{v}_{m,n} \end{pmatrix} \prod_{i=1}^{\nu} (\beta_i - \alpha_i)\mu^{-2\nu},
\end{aligned}
$$

where $\phi(A, B)=\psi[F^{-1}(A, B, 1), B] \cdot J\binom{V}{A}$ and $\bar{u}_{m,n}$, $\bar{v}_{m,n}$ is a properly chosen point of the region $r_{m,n}$. Equation (13) can also be written in the form

$$
\begin{aligned}
&\text{prob } (\alpha \leqq a - [a] < \beta) \\
(14)\qquad &= \prod_{i=1}^{\nu} (\beta_i - \alpha_i) \sum_{m,n} \phi(\overline{A}_{m,n}, \overline{B}_{m,n}) J\begin{pmatrix} A, & B \\ \bar{u}'_{m,n}, & \bar{v}'_{m,n} \end{pmatrix} \mu^{-2\nu}(1 + \epsilon_{m,n}),
\end{aligned}
$$

where $$J\begin{pmatrix} A, & B \\ \bar{u}'_{m,n}, & \bar{v}'_{m,n} \end{pmatrix}\mu^{-2\nu}$$

is the volume of $R_{m,n}$ and $\epsilon_{m,n}$ approaches 0 uniformly with μ. Therefore

$$\lim_{\mu\to 0} \text{prob}\,(\alpha \leqq a - [a] < \beta) = \prod_{i=1}^{\nu}(\beta_i - \alpha_i)\int_\Gamma \phi(A, B)d\sigma \tag{15}$$

$$= \prod_{i=1}^{\nu}(\beta_i - \alpha_i).$$

It follows that

$$\lim_{\mu\to 0} \text{prob}\,(a - [a]\epsilon E) = m(E) \tag{16}$$

for every set E whose frontier points are of measure 0. Thus the distribution of $a-[a]$ approaches uniformity as the parameter μ approaches 0.

We shall consider two simple extensions of this theory. We may be concerned with problems involving only λ of the coordinates of the rest position where $\lambda<\nu$. In this case it is only necessary to observe that any sub-set of the sub-space of λ dimensions is also a sub-set of the space Δ and that if the frontier points of such a set are of measure 0 with respect to the sub-space, they are also of measure 0 with respect to Δ. Hence a uniform distribution is approached in the sub-space.

Suppose that we make a series of n experiments with a mechanical system and that the initial conditions for these experiments are dependent. We shall inquire the limiting value of the probability that the rest positions of the 1st, 2nd, $\cdots$, nth experiments will fall respectively in the sets $E_1, E_2, \cdots, E_n$. This problem is essentially the same as those previously considered. Thus we have a system of $n\nu$ second order differential equations obtained by repeating the ν equations n times. We have a corresponding distribution function for the initial conditions which is defined throughout a phase space of $2n\nu$ dimensions. It follows that the probability in question is $m(E_1)\cdot m(E_2)\cdots m(E_n)$. Hence in the limit we have independence with respect to the rest positions even though we had dependence in the initial conditions.

THE UNIVERSITY OF MICHIGAN

GENERALIZED JACOBI POLYNOMIALS

D. N. SEN AND V. RANGACHARIAR

1. *Introduction.* The differential equation

$$(\alpha x^2 + \beta x + \gamma)\frac{d^2y}{dx^2} - (x + a_1)\frac{dy}{dx} + \{n - n(n-1)\alpha\}y = 0,$$

where n is a positive integer, has polynomial solutions y_n of degree n. Some properties of these polynomials have been obtained by Brenke* and by Lawton.† The object of this paper is to derive fresh properties and in particular to study the zeros of these polynomials. Brenke proved that

$$h_n y_n = \frac{1}{\rho} D^n\{\rho P^n\},$$

where

$$P \equiv \alpha x^2 + \beta x + \gamma \equiv -\alpha(x-a)(b-x), \quad (a < b),$$

$$\rho \equiv -\frac{1}{\alpha}(x-a)^{A-1}(b-x)^{B-1},$$

$$A = \frac{a + a_1}{\alpha(b-a)}, \qquad B = \frac{b + a_1}{-\alpha(b-a)},$$

and h_n is the coefficient of x^n in the right-hand side. It has also been proved by him that if A and B are positive, the following recurrence formula holds good.

(A) $$y_n = (a_n + x)y_{n-1} - b_n y_{n-2},$$

where

$$b_n = \frac{c_n^2}{c_{n-1}^2} \quad \text{and} \quad \frac{1}{c_n^2} = \int_a^b \rho y_n^2\, dx,$$

$$a_n = -c_{n-1}^2 k_{n-1}, \quad \text{where} \quad k_n = \int_a^b x\rho y_n^2\, dx.$$

* This Bulletin, vol. 36 (1930).

† This Bulletin, vol. 38 (1932).

The values of A and B have been assumed to be positive for the existence of the integral $\int_a^b \rho dx$.

2. *Values of y_n at a and b.*

$$h_n y_n = \frac{1}{\rho} D^n(\rho P^n) = \frac{(-\alpha)^{n-1}}{\rho} D^n\left\{(x-a)^{n+A-1}(b-x)^{n+B-1}\right\}$$
$$= (-\alpha)^n (x-a)^{1-A}(b-x)^{1-B} D^n\left\{(x-a)^{n+A-1}(b-x)^{n+B-1}\right\}$$
$$= (-\alpha)^n (x-a)^{1-A}(b-x)^{1-B} \cdot \left[\sum_{r=0}^{+n} (-1)^r C_{n,r} S_r (x-a)^{r+A-1}(b-x)^{n-r+B-1}\right],$$

where

$$S_r = (n+A-1)(n+A-2)\cdots(A+r)$$
$$\times (n+B-1)(n+B-2)\cdots(n+B-r).$$

Now $h_n = \alpha^n(2n+A+B-2)(2n+A+B-3)\cdots(n+A+B-1)$. Hence

$$y_n(a) = (-1)^n \frac{(n+A-1)(n+A-2)\cdots A(b-a)^n}{(2n+A+B-2)(2n+A+B-3)\cdots(n+A+B-1)},$$

$$y_n(b) = \frac{(n+B-1)(n+B-2)\cdots B(b-a)^n}{(2n+A+B-2)(2n+A+B-3)\cdots(n+A+B-1)}.$$

3. *Value of b_n in formula* (A).

$$\int_a^b \rho y_k^2 dx = \frac{1}{h_k^2}\int_a^b D^k(\rho P^k)\left\{\frac{1}{\rho} D^k(\rho P^k)\right\} dx$$
$$= \frac{(-1)^k}{h_k^2}\int_a^b \rho P^k D^k\left\{\frac{1}{\rho} D^k(\rho P^k)\right\} dx$$
$$= \frac{(-1)^k k!}{h_k}\int_a^b \rho P^k dx$$
$$= \frac{(-1)^k k!}{h_k}(-\alpha)^{k-1}\int_a^b (x-a)^{k+A-1}(b-x)^{k+B-1} dx$$
$$= \frac{(-1)^k k!}{h_k}(-\alpha)^{k-1}\frac{\Gamma(k+A)\Gamma(k+B)}{\Gamma(2k+A+B)}(b-a)^{2k+A+B-1}.$$

Hence

$$b_n = \frac{\int_a^b \rho y_{n-1}^2 dx}{\int_a^b \rho y_{n-2}^2 dx}$$

$$= \alpha \frac{h_{n-2}}{h_{n-1}} (n-1) \frac{(n+A-2)(n+B-2)}{(2n+A+B-3)(2n+A+B-4)} (b-a)^2$$

$$= (n-1) \frac{(n+A-2)(n+B-2)(n+A+B-3)}{(2n+A+B-3)(2n+A+B-4)^2(2n+A+B-5)} (b-a)^2.$$

4. *Relation among y_n', y_n, y_{n+1}.* From §2

$$h_n y_n = (-\alpha)^n \sum_{r=0}^{r=n} (-1)^r S_r C_{n,r} (x-a)^r (b-x)^{n-r}.$$

Hence

$$h_n y_n' = (-\alpha)^n \sum_{0}^{n} (-1)^r S_r C_{n,r} \{ r(x-a)^{r-1}(b-x)^{n-r} - (n-r)(x-a)^r (b-x)^{n-r-1} \}.$$

The coefficient

$$(x-a)^r (b-x)^{n-r-1}$$

on the right-hand side under summation is

$$(-1)^{r+1} C_{n,r+1}(r+1) S_{r+1} + (-1)^{r+1}(n-r) C_{n,r} S_r$$

$$= (-1)^{r+1} n C_{n-1,r} \{ S_r + S_{r+1} \}$$

$$= (-1)^{r+1} n C_{n-1,r} \{ (A+r) + (n+B-r-1) \} (n+A-1) \cdots (A+r+1) \times (n+B-1) \cdots (n+B-r)$$

$$= (-1)^{r+1} n C_{n-1,r} (n+A+B-1)(n+A-1) \cdots (A+r+1) \times (n+B-1) \cdots (n+B-r).$$

Therefore

$$h_n y_n' = -(-\alpha)^n n(n+A+B-1) \times \sum_{r=0}^{n-1} \{ (-1)^r C_{n-1,r} (n+A-1) \cdots (A+r+1) \times (n+B-1) \cdots (n+B-r) \times (x-a)^r (b-x)^{n-r-1} \}.$$

Now

$$\{(n+A)(b-x)-(n+B)(x-a)\}h_n y_n$$
$$=\{(n+A)(b-x)-(n+B)(x-a)\}(-\alpha)^n \cdot \sum_{r=0}^{n}(-1)^r S_r C_{n,r}(x-a)^r(b-x)^{n-r}.$$

The coefficient of $(x-a)^{r+1}(b-x)^{n-r}$ on the right-hand side is

$$(-\alpha)^n(-1)^{r+1}\{(n+B)C_{n,r}S_r+(n+A)C_{n,r+1}S_{r+1}\}$$
$$=(-\alpha)^n(-1)^{r+1}\{(n+B)(A+r)C_{n,r}+(n+A)(n+B-r-1)C_{n,r+1}\}$$
$$\times(n+A-1)\cdots(A+r+1)$$
$$\times(n+B-1)\cdots(n+B-r).$$

The expression within the braces is

$$(n+B)(n+A-\overline{n-r})C_{n,r}+(n+A)(n+B-\overline{r+1})C_{n,r+1}$$
$$=(n+A)(n+B)(C_{n,r}+C_{n,r+1})-nC_{n-1,r}(2n+A+B)$$
$$=(n+A)(n+B)C_{n+1,r+1}-nC_{n-1,r}(2n+A+B).$$

Hence

$$\{(n+A)(b-x)-(n+B)(x-a)\}h_n y_n$$
$$=(-\alpha)^n\sum_{0}^{n+1}(-1)^r C_{n+1,r}(n+A)(n+A-1)\cdots(A+r)$$
$$\times(n+B)\cdots(n+B-r+1)\times(x-a)^r(b-x)^{n-r+1}$$
$$+(-\alpha)^n n(2n+A+B)\cdot\sum_{0}^{n-1}(-1)^r C_{n-1,r}(n+A-1)\cdots(A+r+1)$$
$$\times(n+B-1)\cdots(n+B-r)(x-a)^{r+1}(b-x)^{n-r}$$
$$=\frac{h_{n+1}}{-\alpha}y_{n+1}+(-\alpha)^n n(2n+A+B)\cdot\sum_{0}^{n-1}(-1)^r C_{n-1,r}(n+A-1)\cdots(A+r+1)$$
$$\times(n+B-1)\cdots(n+B-r)(x-a)^{r+1}(b-x)^{n-r}.$$

Also

$$(x-a)(b-x)h_n y_n' = -n(n+A+B-1)(-\alpha)^n \cdot \sum_0^{n-1} (-1)^r C_{n-1,r}(n+A-1)\cdots(A+r+1) \times (n+B-1)\cdots(n+B-r)(x-a)^{r+1}(b-x)^{n-r}.$$

Hence

$$\{(n+A)(b-x)-(n+B)(x-a)\}h_n y_n = \frac{h_{n+1}}{-\alpha} y_{n+1} - \frac{2n+A+B}{n+A+B-1}(x-a)(b-x)h_n y_n'.$$

Therefore

$$(x-a)(b-x)y_n' = -(2n+A+B-1)y_{n+1} + \frac{n+A+B-1}{2n+A+B}\{(n+B)(x-a)-(n+A)(b-x)\}y_n.$$

5. *Results Valid for A and B Negative.* If A and B be assumed to be negative, the results of §§2 and 4 obviously continue to hold good. The recurrence formula (A), being an identity and y_n''s being polynomials, will still be true and the form for b_n obtained in §3 will also hold good.

6. *Roots of* $y_n=0$. Lawton and Fujiwara proved that if A and B be negative or zero, and p, q positive integers such that $0<A+p\leqq 1$, $0\leqq B+q\leqq 1$, the number of roots of $y_n=0$ inside (a, b) is $n-p-q$, $(n\geqq p+q+1)$. We shall derive the result by arguments based on methods of Sturm's theorem in the theory of equations and get fresh results by this consideration. The set of functions $y_n, y_{n-1}, \cdots, y_{p+q}$ is taken for this purpose, it being proved in the next section that $y_{p+q}=0$ has no roots inside (a, b).

7. *No roots of* $y_{p+q}=0$ *inside* (a, b). When x lies between a and b, $(x-a)$ and $(b-x)$ are both positive. The coefficient of $(x-a)^r(b-x)^{n-r}$ on the right-hand side of $h_n y_n$ in §2 is $(-\alpha)^n C_{n,r}(-1)^r S_r$. When $n=p+q$, this coefficient becomes

$$(-\alpha)^{p+q}C_{p+q,r}(-1)^r(p+q+A-1)\cdots(A+r) \times (p+q+B-1)\cdots\times(p+q+B-r).$$

If $r=p-s$, where s is positive or zero,

$$(p+q+A-1)\cdots(p+A-1)\cdots(A+p-s)$$

has s negative factors and in $(p+q+B-1)\cdots(q+B+s)$ all the factors are positive. Hence in this case the sign of $(-1)^r S_r$ is that of $(-1)^p$. If, however, $r=p+t$, where t is positive, all the factors in $(p+q+A-1)\cdots(A+p+t)$ are positive and in $(p+q+B-1)\cdots(q+B-1)\cdots(q+B-t)$ there are t negative factors. Hence again the sign of $(-1)^r S_r$ is that of $(-1)^p$. It follows that $h_{p+q}y_{p+q}$ cannot vanish within (a, b) and has the same sign as that of $(-\alpha)^{p+q}(-1)^p$. Now

$$h_{p+q} = (+\alpha)^{p+q}(2p+2q+A+B-2)\cdots(p+q+A+B)(p+q+A+B-1).$$

Hence the sign of y_{p+q} is that of $(-1)^q(p+q+A+B-1)$, that is, the sign of y_{p+q} within (a, b) is that of $(-1)^q$ or $(-1)^{q-1}$ according as the sum of the fractional parts of $-A$ and $-B$ is <1 or >1. The case in which the sum $=1$ follows by continuity.*

8. *Changes in Sign.* In passing through a root θ of $y_n=0$ a change of sign will be lost or gained between y_n and y_{n-1} according as $y_n'(\theta)$ and $y_{n-1}(\theta)$ have like or unlike signs, that is, according as b_{n+1} is positive or negative (§4 and (A)). By formula (A) when $y_{k-1}=0$, y_k/y_{k-2} will have sign opposite to that of b_k. If b_k be positive, no change of sign will be lost or gained in passing through a root δ of $y_{k-1}=0$; if on the other hand b_k be negative, two changes of sign will be lost or gained according as $y_{k-1}'(\delta)$ and $y_{k-2}(\delta)$ are of the same or opposite sign. The relation of §4 enables us to deal with the different possibilities and an examination of the changes of sign lost enables us to determine the number of roots of $y_n=0$ within (a, b).

9. *Sign of b_n for Different Values of n.* For studying the zeros of y_n, $(n \geqq p+q)$, it is enough to consider the sign of b_n for $n \geqq p+q+2$. All the factors in b_n except $n+A+B-3$ are evidently positive. This factor is also positive for $n \geqq p+q+2$ if the sum of the fractional parts of $-A$ and $-B$ be less than unity, but if this sum exceeds unity, then b_n is negative for $n=p+q+2$, and positive for subsequent values of n.

* This re-statement is made to guard against the possibility that we may have $h_{p+q}=0$ in the fraction $(-\alpha)^{p+q}(-1)^p/h_{p+q}$.

10. *Sum of Fractional Parts* <1. Thus if the sum of the fractional parts be <1, a change of sign will be lost in passing through a root of $y_n=0$ and no change of sign will be lost or gained for a root of any of the intermediate functions. Also from §2, it is evident that at b the functions are of the same sign and at a consecutive ones are of opposite signs. Hence the number of roots of $y_n=0$ between a and b will be $n-p-q$, $(n\geqq p+q)$.

11. *Sum of Fractional Parts* >1. If, on the other hand, the sum of the fractional parts >1, at $x=b$ the signs of y_{p+q} and y_{p+q+1} are opposite and there will be continuation of sign for the subsequent functions; whereas at a, y_{p+q} and y_{p+q+1} will be of the same sign and then there will be variation of sign at each stage subsequently. Hence the number of changes of sign lost will be $n-p-q-2$. The equation $y_{p+q+1}=0$ will evidently have only one root in (a, b). In passing through this root there will be two gains as b_{p+q+2} is negative in this case. In passing through a root of $y_n=0$, $(n\geqq p+q+2)$, there will be a loss of sign between y_n and y_{n-1} and there will be no loss or gain for a root of any $y_k=0$, $(p+q+2\leqq k\leqq n-1)$. Hence the number of roots of $y_n=0$ will exceed by 2 the number of changes of sign lost, that is, it will be $n-p-q$, $(n\geqq p+q)$. The result of Lawton is extended by our method to the case $n=p+q$.

12. *Application of Sturm's Method.* Sturm's method may be applied for the ranges $(-\infty, a)$ and (b, ∞) as well. In particular we shall show that the number of roots of the equation $y_{n''}=0$ and $y_{n'}=0$, where $n''>n'\geqq[-(A+B)+3]=n_0$, is the same in either of these ranges, $[x]$ denoting the integral part of x. For studying such a range it is clear that we cannot stop at y_{p+q}, but we have to take the whole set of functions down to y_0. Let us take the sets $y_0, y_1, \cdots, y_{n'}$ and $y_0, y_1, \cdots, y_{n''}$. In the first instance, the number of changes of sign in passing through $(-\infty, a)$ will be the same for the two sets. Also b_{n+1} is positive for $n\geqq n_0$. Hence in the second set no change is contributed for passing through a root of $y_{n'}=0, y_{n'+1}=0, \cdots, y_{n''-1}=0$, which come as intermediate functions. Again the nature of the change for a root of the last function in the two cases will be the same, for $y'_{n'}(\theta')/y_{n'-1}(\theta')$ and $y'_{n''}(\theta'')/y_{n''-1}(\theta'')$ are of the same sign $[y'_n(\theta)/y_{n'-1}(\theta)$ is negative in $(-\infty, a)$ for $n\geqq n_0]$. Now in the first set the number of losses is due to (1) the roots of some inter-

mediate functions and (2) the roots of $y_{n'}=0$. In the second set the contribution due to (1) persists. By the introduction of the additional functions in the second set, the change contributed by a root of $y_{n'}=0$ in the first set is transferred to a root of $y_{n''}=0$ in the second set. Also no change arises for a root of $y_{n'}=0$, $y_{n'+1}=0, \cdots, y_{n''-1}=0$. Since the total number of losses is the same in the two cases, the conclusion will be that the number of roots for $y_{n'}=0$ and $y_{n''}=0$ in the range $(-\infty, a)$ will be the same; similarly for the range (b, ∞). Also in (a, b) the numbers of roots of $y_{n'}=0$ and $y_{n''}=0$ are $n'-p-q$ and $n''-p-q$. Hence we conclude that the number of imaginary roots of $y_{n'}=0$ and $y_{n''}=0$ will be the same for $n''>n'\geqq n_0$.

SCIENCE COLLEGE,
PATNA, INDIA

NOTE ON THE EXISTENCE OF AN nTH DERIVATIVE DEFINED BY MEANS OF A SINGLE LIMIT

BY NORMAN MILLER

The nth derivative of a function $f(x)$ may be defined without the use of derivatives of lower order by means of the limit of a certain quotient. Conditions necessary and sufficient for the existence and continuity of $f^{(n)}(x)$ at a point $x=a$ and also for the mere existence of $f^{(n)}(a)$ have been recently given by Franklin.* The purpose of the present note is to state necessary and sufficient conditions of a somewhat more general form with proofs which use only Rolle's theorem and elementary properties of determinants.

Let $f_i(x)$ and $\phi_i(x)$, $(i=1, 2, \cdots, n+1)$, be functions possessing derivatives of the nth order, continuous in an interval I. Let $x_1, x_2, \cdots, x_{n+1}$ be points of I which close down in an arbitrary manner on a point a, in the sense that

$$(1) \qquad |x_j - a| < \epsilon_k, \qquad \lim_{k\to\infty} \epsilon_k = 0.$$

We shall use the notation

* This Bulletin, vol. 41 (1935), p. 573.

$$(2) \qquad B_{n,k} = \frac{\begin{vmatrix} f_1(x_j) & f_2(x_j) & \cdots & f_{n+1}(x_j) \end{vmatrix}}{\begin{vmatrix} \phi_1(x_j) & \phi_2(x_j) & \cdots & \phi_{n+1}(x_j) \end{vmatrix}},$$

in which numerator and denominator are determinants of which only the jth row of each is written. The following theorem* requires for its proof only a repeated use of Rolle's theorem

$$(3) \qquad B_{n,k} = \frac{\begin{vmatrix} f_1(\xi_1) & \cdots & f_{n+1}(\xi_1) \\ f_1'(\xi_2) & \cdots & f_{n+1}'(\xi_2) \\ \cdots & \cdots & \cdots \\ f_1^{(n)}(\xi_{n+1}) & \cdots & f_{n+1}^{(n)}(\xi_{n+1}) \end{vmatrix}}{\begin{vmatrix} \phi_1(\xi_1) & \cdots & \phi_{n+1}(\xi_1) \\ \phi_1'(\xi_2) & \cdots & \phi_{n+1}'(\xi_2) \\ \cdots & \cdots & \cdots \\ \phi_1^{(n)}(\xi_{n+1}) & \cdots & \phi_{n+1}^{(n)}(\xi_{n+1}) \end{vmatrix}},$$

where $\xi_1 = x_1$ and $\xi_2, \cdots, \xi_{n+1}$ lie between the extreme points of x_j. If the points x_j close down in any manner on a point a of I, then $\xi_j \to a$, $(j=1, 2, \cdots, n+1)$, and on account of the continuity of the derivatives we have the following theorem.

THEOREM 1. *If the points x_j close down on a in an arbitrary manner subject to* (1) *and $W(\phi_1, \phi_2, \cdots, \phi_{n+1}) \neq 0$ at the point a, where W denotes the Wronskian of the functions indicated, then $B_{n,k}$ has a unique limit*

$$(4) \qquad \lim_{k \to \infty} B_{n,k} = \left[\frac{W(f_1, \cdots, f_{n+1})}{W(\phi_1, \cdots, \phi_{n+1})}\right]_{x=a}.$$

Conversely, however, it is not true that if $B_{n,k}$ has a unique limit for any approach of the points to a, then the derivatives involved in the Wronskians all exist. If, for example, $f_j \equiv \phi_j$, $(j=1, 2, \cdots, n+1)$, then $B_{n,k} \equiv 1$ and $\lim B_{n,k} = 1$ for any ap-

* Pólya and Szegö, *Aufgaben und Lehrsätze*, vol. 2, p. 54. The theorem in a slightly less general form was given by Schwarz, Annali di Matematica, (2), vol. 10 (1880), p. 129; Collected Works, vol. 2, p. 296. A proof based on Rolle's theorem was given by Stieltjes, Nouvelles Annales de Mathématique, (3), vol. 7 (1888), p. 26; Collected Works, p. 105. See also Miller, Transactions of the Royal Society of Canada, 1931, p. 195.

proach, even though the functions be non-differentiable. If, however, all the functions but one, say $f_1(x)$, possess nth derivatives continuous at a and are subject to certain further restrictions stated below, the existence of a unique limit for $B_{n,k}$ as the points x_j close down on a in an arbitrary manner* furnishes a sufficient condition for the existence and continuity of $f_1^{(n)}(x)$ at $x=a$. If the restriction be made that one of the points x_j remain at a while the others approach a in any manner, the existence of $\lim B_{n,k}$ in this case suffices to ensure the existence of $f_1^{(n)}(a)$ without the continuity of the derivative at this point. This latter theorem will first be proved.

THEOREM 2. *Hypotheses*: (i) *In an interval I containing the point a the n functions f_i, $(i=2, 3, \cdots, n+1)$, are linearly independent in every subinterval and possess derivatives of the nth order*, (ii) *the functions ϕ_i, $(i=1, 2, 3, \cdots, n+1)$, possess derivatives of the nth order*, (iii) *$B_{n,k}$ possesses a unique limit as the point x_1 remains at a while the points $x_2, \cdots, x_{n+1}$ close down on a in an arbitrary manner*, (iv) *$W(f_2, f_3, \cdots, f_{n+1}) \neq 0$ at the point a. Then f_1 also possesses a derivative of the nth order at $x=a$ and the derivatives of f_1 satisfy the relation* (4).

We prove first the existence of $f_1'(a)$. For this purpose set $x_1=a$ and replace the elements of the second row of the numerator of (2) by $[f_j(x_2)-f_j(a)]/(x_2-a)$, $(j=1, 2, \cdots, n+1)$, with a similar replacement in the second row of the denominator. Now, leaving $x_3, \cdots, x_{n+1}$ fixed, let $x_2 \to a$. By hypothesis

* Since the denominator of $B_{n,k}$ vanishes whenever two or more of the points x_j coincide, the quotient is not defined at such points. The hypothesis here stated shall be understood to require that the discontinuities at all such points be removable. That is, if in a sufficiently small neighbourhood of a some of the points x_j remain fixed while the remainder approach a point (either a or a_1 near a) in an arbitrary manner, the limit of $B_{n,k}$ must exist. The following example is instructive. On the parabola $y=x^2$ take a set of points whose abscissas are 1, 1/2, 1/4, 1/8, $\cdots$. Join successive points of this sequence by chords, forming a broken line curve, to which the limiting point (0, 0) is adjoined. Call $f(x)$ the function represented by this broken line curve. It has the property that $\lim_{x_1 \to 0,\ x_2 \to 0} [f(x_2)-f(x_1)]/(x_2-x_1)=0$ for $x_2 \neq x_1$, also $\lim_{x_1 \to 0} \overline{\lim}_{x_2 \neq x_1} [f(x_2)-f(x_1)]/(x_2-x_1)=0$ and $\lim_{x_1 \to 0} \underline{\lim}_{x_2 \neq x_1} [f(x_2)-f(x_1)]/(x_2-x_1)=0$. But $\lim_{x_1 \to 0,\ x_2 \to 0} [f(x_2)-f(x_1)]/(x_2-x_1)$ does not exist in the sense in which we are using the term, since the quotient has irremovable discontinuities at a set of points condensing on the origin.

(iii) $B_{n,k}$ must then approach a limit (if $x_3, \cdots, x_{n+1}$ are close enough to a). Also $\lim [f_j(x_2)-f_j(a)]/(x_2-a)=f_j'(a)$ for $j\neq 1$ and $\lim [\phi_j(x_2)-\phi_j(a)]/(x_2-a)=\phi_j'(a)$ for $j=1, 2, \cdots, n+1$. The only element whose limit is in doubt is $[f_1(x_2)-f_1(a)]/(x_2-a)$. It will follow that this element approaches a limit and therefore that $f_1'(a)$ exists as soon as we show that its cofactor in the numerator is not zero. This cofactor is, except for sign,

$$C_{12}=\begin{vmatrix} f_2(a) & f_3(a) & \cdots & f_{n+1}(a) \\ f_2(x_3) & f_3(x_3) & \cdots & f_{n+1}(x_3) \\ f_2(x_4) & f_3(x_4) & \cdots & f_{n+1}(x_4) \\ \cdots & \cdots & \cdots & \cdots \\ f_2(x_{n+1}) & f_3(x_{n+1}) & \cdots & f_{n+1}(x_{n+1}) \end{vmatrix}.$$

The elements of the first row are not all zero, by hypothesis (iv). In any subinterval of I a point x_3 exists such that not all the two-rowed minors of the first two rows are zero, for if not the n functions would be linearly dependent in this subinterval. Let such a point x_3 be chosen. Then in any subinterval of I, x_4 may be chosen such that not all the three-rowed minors of the first three rows are zero, and so on. Finally, $x_3, x_4, \cdots, x_{n+1}$ may be chosen in an infinite variety of ways such that $C_{12}\neq 0$. The existence of $f_1'(a)$ is then established.

Proceeding now by induction suppose that $f_1^{(r)}(a)$ exists. In order to show the existence of $f_1^{(r+1)}(a)$ we first transform $B_{n,k}$ in the manner in which (3) is obtained but without carrying the transformation so far. Consider

$$\begin{vmatrix} f_1(x_1) & f_2(x_1) & \cdots & f_{n+1}(x_1) \\ f_1(x_2) & f_2(x_2) & \cdots & f_{n+1}(x_2) \\ \cdots & \cdots & \cdots & \cdots \\ f_1(x_{n+1}) & f_2(x_{n+1}) & \cdots & f_{n+1}(x_{n+1}) \end{vmatrix} - B_{n,k}\begin{vmatrix} \phi_1(x_1) & \phi_2(x_1) & \cdots & \phi_{n+1}(x_1) \\ \phi_1(x_2) & \phi_2(x_2) & \cdots & \phi_{n+1}(x_2) \\ \cdots & \cdots & \cdots & \cdots \\ \phi_1(x_{n+1}) & \phi_2(x_{n+1}) & \cdots & \phi_{n+1}(x_{n+1}) \end{vmatrix}. \tag{5}$$

Replace x_{n+1} by x. The resulting function of x vanishes at

$x=x_{n+1}$ and at* $x=x_n$. Hence its derivative vanishes at x_{n+1}, between x_n and x_{n+1}. In the last rows of (5) we may replace $f_j(x_{n+1})$ by $f_j'(x_{n+1}')$ and $\phi_j(x_{n+1})$ by $\phi_j'(x_{n+1}')$, so that the resulting expression vanishes. By a similar operation we may replace the elements of the second last rows by derivatives and x_n by x_n' which is between x_{n-1} and x_n. This may be repeated until all the elements except those of the first rows have been replaced by derivatives, the resulting expression being zero. Beginning again at the last rows replace x_{n+1}' by x. The resulting function of x vanishes at x_{n+1}' and at x_n'. An application of Rolle's theorem leads to the substitution of second derivatives in the last rows, and of x_{n+1}'' for x_{n+1}', where x_{n+1}'' is between x_n' and x_{n+1}'. Repetitions of the operation will replace by second derivatives all the elements of the rows beginning with the third. In a similar manner we replace by third derivatives all the elements of the rows beginning with the fourth. Continuing this process until we arrive at rth derivatives, we get for $B_{n,k}$ a quotient of which the numerator is

$$(6)\qquad \begin{vmatrix} f_1(\eta_1) & f_2(\eta_1) & \cdots & f_{n+1}(\eta_1) \\ f_1'(\eta_2) & f_2'(\eta_2) & \cdots & f_{n+1}'(\eta_2) \\ \cdots & \cdots & \cdots & \cdots \\ f_1^{(r)}(\eta_{r+1}) & f_2^{(r)}(\eta_{r+1}) & \cdots & f_{n+1}^{(r)}(\eta_{r+1}) \\ f_1^{(r)}(\eta_{r+2}) & f_2^{(r)}(\eta_{r+2}) & \cdots & f_{n+1}^{(r)}(\eta_{r+2}) \\ \cdots & \cdots & \cdots & \cdots \\ f_1^{(r)}(\eta_{n+1}) & f_2^{(r)}(\eta_{n+1}) & \cdots & f_{n+1}^{(r)}(\eta_{n+1}) \end{vmatrix},$$

and the denominator is a similar expression with ϕ replacing f, where $\eta_1=x_1$, η_2 is between x_1 and x_2, η_3 is between the extremes of x_1, x_2 and x_3, and so on.

We now make use of the resulting expression for $B_{n,k}$ as follows. Let $x_1=x_2=\cdots=x_{r+1}=a$. Then $\eta_1=\eta_2=\cdots=\eta_{r+1}=a$. In (6) replace $f_j^{(r)}(\eta_{r+2})$ by

$$\frac{f_j^{(r)}(\eta_{r+2})-f_j^{(r)}(a)}{\eta_{r+2}-a}, \qquad (j=1, 2, \cdots, n+1),$$

* For the deduction of the result which follows, the points x_j are taken distinct. When, in the result, some of the points are made to coincide, it is the limit of $B_{n,k}$ which is used.

with similar replacements in the $(r+2)$th row of the denominator of $B_{n,k}$. Thus (6) becomes

$$(7)\qquad \begin{vmatrix} f_1(a) & f_2(a) & \cdots & f_{n+1}(a) \\ f_1'(a) & f_2'(a) & \cdots & f_{n+1}'(a) \\ \cdots & \cdots & \cdots & \cdots \\ f_1^{(r)}(a) & f_2^{(r)}(a) & \cdots & f_{n+1}^{(r)}(a) \\ \dfrac{f_1^{(r)}(\eta_{r+2}) - f_1^{(r)}(a)}{\eta_{r+2} - a} & & \cdots & \dfrac{f_{n+1}^{(r)}(\eta_{r+2}) - f_{n+1}^{(r)}(a)}{\eta_{r+2} - a} \\ f_1^{(r)}(\eta_{r+3}) & & \cdots & f_{n+1}^{(r)}(\eta_{r+3}) \\ \cdots & \cdots & \cdots & \cdots \\ f_1^{(r)}(\eta_{n+1}) & & \cdots & f_{n+1}^{(r)}(\eta_{n+1}) \end{vmatrix}.$$

Let $x_{r+2} \to a$. Then $\eta_{r+2} \to a$. The ratio of which (7) is the numerator approaches a limit by hypothesis (iii). The only element in numerator or denominator whose limit is in doubt is $[f_1^{(r)}(\eta_{r+2}) - f_1^{(r)}(a)]/(\eta_{r+2} - a)$. Consider the cofactor of this element in (7),

$$(8)\qquad \begin{vmatrix} f_2(a) & \cdots & f_{n+1}(a) \\ f_2'(a) & \cdots & f_{n+1}'(a) \\ \cdots & \cdots & \cdots \\ f_2^{(r)}(a) & \cdots & f_{n+1}^{(r)}(a) \\ f_2^{(r)}(\eta_{r+3}) & \cdots & f_{n+1}^{(r)}(\eta_{r+3}) \\ \cdots & \cdots & \cdots \\ f_2^{(r)}(\eta_{n+1}) & \cdots & f_{n+1}^{(r)}(\eta_{n+1}) \end{vmatrix}.$$

Not all the $(r+1)$-rowed minors of the first $r+1$ rows of (8) vanish, by hypothesis (iv). If it should happen that all the $(r+2)$-rowed minors of the first $r+2$ rows are zero, then a slight change in x_{r+3} will cause a change in η_{r+3} and will lead to $(r+2)$-rowed minors not all of which are zero, on account of hypothesis (i). Similarly, $x_{r+3}, \cdots, x_{n+1}$ may be chosen in an infinite variety of ways so that (8) is not zero. Hence $\lim_{\eta_{r+2} \to a} [f_1^{(r)}(\eta_{r+2}) - f_1^{(r)}(a)]/(\eta_{r+2} - a)$ or $f_1^{(r+1)}(a)$ exists. This completes the induction and establishes Theorem 2.

In order that the derivative $f^{(n)}(x)$ be continuous at $x = a$ the hypotheses must be made slightly more stringent. The theorem

is suggested by the corresponding theorem of Franklin already referred to and the proof makes use of a device which he introduced.

THEOREM 3. *The hypotheses of Theorem 2 are postulated, with the following modifications. In* (i) *and* (ii) *the nth derivatives of* $f_2, \cdots, f_{n+1}, \phi_1, \cdots, \phi_{n+1}$ *shall be continuous.* (iii) *becomes:* $B_{n,k}$ *possesses a unique limit as the points* $x_1, x_2, \cdots, x_{n+1}$ *close down on* a *in an arbitrary manner. The conclusion is that* $f_1^{(n)}(x)$ *exists and is continuous at* $x=a$.

For the proof the existence of $f_1'(a)$ follows from Theorem 2. Take a_1, a point neighboring to a. Let $x_1=a_1$ and let $x_2, \cdots, x_{n+1}$ approach a_1 and then let a_1 approach a. Since $B_{n,k}$ has the same limit for every approach of the points to a, and since the derivatives of the functions other than f_1 are continuous, it follows that $\lim_{a_1 \to a} \lim_{x_2 \to a_1} [f_1(x_2)-f_1(a_1)]/(x_2-a_1) = f_1'(a)$, which proves the continuity of $f_1'(x)$ at a. In the same way the existence proof for $f_1^{(r+1)}(a)$ may be modified, with the hypotheses now available, to include the proof of the continuity of this derivative. Hence the induction for this case is also complete.

PARTICULAR CASES. Set $f_2(x)=x^{n-1}, \cdots, f_n(x)=x, f_{n+1}(x)=1$; $\phi_1(x)=x^n, \phi_2(x)=x^{n-1}, \cdots, \phi_{n+1}(x)=1$ and denote $B_{n,k}$ for this case by $A_{n,k}$. Under the hypothesis of a unique limit for $A_{n,k}$ we have,* dropping the subscript of $f_2(x)$, $\lim_{k\to\infty} A_{n,k} = f^{(n)}(a)/n!$ and $f^{(n)}(a) = n! \lim_{k\to\infty} A_{n,k}$. Here the mere existence of $f^{(n)}(a)$ follows or the existence and continuity of $f^{(n)}(x)$ at $x=a$ according as the method of approach of the points to a is that specified in Theorem 2 or in Theorem 3.

Again set $f_2(x)=x^{n-1}, \cdots, f_n(x)=x, f_{n+1}(x)=1$. Under the pertinent hypotheses the expression for $f^{(n)}(a)$ is, in this case,

$$f^{(n)}(a) = (-1)^n \frac{W(\phi_1, \cdots, \phi_{n+1})}{(n-1)!(n-2)! \cdots 2!1!} \lim_{k\to\infty} B_{n,k},$$

the value of the Wronskian being taken at the point a.

QUEEN'S UNIVERSITY,
KINGSTON, CANADA

* Franklin, loc. cit.

A SPECIAL PARAMETERIZATION OF CURVES

BY MARSTON MORSE

1. *Introduction.* Parameterization of curves by means of arc length fails when the arc length is infinite. The present paper develops the properties of a special parameterization of curves which never fails to exist and which is most useful in applications. The special parameter is called a μ-length and is an extension of a function of sets defined by H. Whitney* and applied by Whitney to families of simple non-intersecting curves. The curves employed in the present paper are general continuous images of a line segment. This necessitates a slight modification of the definition of Whitney taking order into account. The properties of μ-length developed here are directed largely towards applications in abstract variational theory. While many of Whitney's ideas go over, there are nevertheless certain sharp differences both in the proofs and in the results. In particular the variational theory of general families of closed curves requires concepts not readily suggested either by the functions of Whitney or the extension developed here. The results of the present paper will be applied in the author's paper entitled *Abstract variational theory*.

2. *The μ-Length.* Let N be a space of points p, q, r with a metric which is not in general symmetric. That is, to each ordered pair p, q of points of N there shall correspond a number denoted by pq such that $pp=0$, $pq>0$ if $p\neq q$, and

$$pq \leqq pr + rq. \tag{1}$$

We term pq the *distance* from p to q. Let $|pq|$ denote the maximum of pq and qp. We term $|pq|$ the *absolute distance* between p and q. We see that $|pq| = |qp|$ and that $|pq| \leqq |pr| + |rq|$. The points of N taken with the distance $|pq|$ form a metric space which we denote by $|N|$. We shall use the metric of N in defining μ-lengths. For other purposes, in particular in defin-

* H. Whitney, *Regular families of curves*, Annals of Mathematics, vol. 34 (1933), pp. 244–270. Also Proceedings of the National Academy of Sciences, vol. 18 (1932), pp. 275–278 and pp. 340–342.

ing continuity on N and $|N|$, we shall use the metric of $|N|$.

Let t be a number on a closed interval* $(0, a)$. Let $f(p, t)$ be a single-valued (numerical) function of p and t for p on $|N|$ and t on $(0, a)$. The function $f(p, t)$ will be termed *continuous* at (q, τ) if $f(p, t)$ tends to $f(q, \tau)$ as a limit as $|pq| + |t-\tau|$ tends to 0. For each value of t on $(0, a)$ and point p on $|N|$ let $\phi(p, t)$ be a point on $|N|$. The (point) function $\phi(p, t)$ will be termed continuous at (q, τ) if the distance $|\phi(p, t)\phi(q, \tau)|$ tends to 0 as a limit as $|pq| + |t-\tau|$ tends to 0.

Let $p(t)$ be a continuous point function of t for t on $(0, a)$. We term $p = p(t)$ a *parameterized curve* λ (written p-curve) and regard two parameterized curves as *identical* if they are defined by the same point function $p(t)$. We also say that λ is the continuous image on $|N|$ of $(0, a)$. In general curves on $|N|$ will be denoted by Greek letters $\alpha, \beta, \cdots$, while points on $|N|$ will be denoted by letters $p, q, r, \cdots$.

Let λ be a p-curve on $|N|$ given as the continuous image $p = p(t)$ of an interval $0 \leqq t \leqq a$. For $n \geqq 2$ let $t_1 \leqq t_2 \leqq \cdots \leqq t_n$ be a set of n values of t on $(0, a)$ and let $(p_1, \cdots, p_n) = S_n$ be the set of corresponding points p on λ. We term S_n an *admissible* set of n points on λ. Let the minimum of the numbers $p_i p_{i+1}$ as i ranges from 1 to $n-1$ be denoted by $d(S_n)$ and let the least upper bound of all such numbers $d(S_n)$ for a fixed n be denoted by $\mu_n(\lambda)$. Following Whitney we then set

$$\mu_\lambda = \frac{\mu_2(\lambda)}{2} + \frac{\mu_3(\lambda)}{4} + \frac{\mu_4(\lambda)}{8} + \cdots . \tag{2}$$

We term the μ_λ the *μ-length* of λ.

We enumerate certain properties of $\mu_n(\lambda)$ and μ_λ.

(a) $\mu_2(\lambda) = d$, *the diameter of* λ.

(b) $\mu_n(\lambda) \leqq d$, *and* $\mu(\lambda) \leqq d$.

(c) $\mu_n(\lambda)$ *tends to* 0 *as* n *becomes infinite.*

---

* For the sake of simplicity we assume that $a > 0$ and that our p-curves do not reduce to points. One could however admit p-curves which reduce to points. The μ-lengths of such curves is zero and we would thus admit intervals $(0, a)$ for which $a = 0$. For such exceptional curves Fréchet distance is defined in the obvious manner. Theorem 2 is obvious if ζ reduces to a point. In the proof of Theorem 4, in case λ_0 reduces to a point, (17) is an easy consequence of $\lambda\lambda_0 < \delta$ provided δ is sufficiently small. Otherwise the theorems and proofs hold as written even when the p-curves reduce to points.

(d) $\mu_n(\lambda) \geqq \mu_{n+1}(\lambda)$.

(e) *If $p(t)$ is not identically constant none of the numbers $\mu_n(\lambda)$ is* 0.

Statements (a), (b), (c), and (e) are obvious. To establish (d) let S_{n+1} be any admissible set of $n+1$ points $p_1, \cdots, p_{n+1}$ on λ, $(n>1)$. There is an integer k such that $p_k p_{k+1} = d(S_{n+1})$. We shall form an admissible set S_n on λ by removing one point from S_{n+1}. If $k \neq 1$ we remove p_1. If $k=1$ we remove p_{n+1}. In both cases the pair p_k, p_{k+1} remains and $d(S_n) = d(S_{n+1})$. The set of all numbers $d(S_{n+1})$ is thus a subset of the numbers $d(S_n)$. Hence (d) is true.

Let $\mu(t)$ be the μ-length of the p-curve on λ defined by $p = p(\tau)$ for τ on the interval $(0, t)$. We shall show that $\mu(t)$ has the following properties.

(f) *$\mu(t)$ is a continuous, non-decreasing function of t.*

(g) *$\mu(t)$ is constant on each interval on which $p(t)$ is constant.*

(h) *$\mu(t)$ is constant on no interval on which $p(t)$ is not constant.*

It is clear that $\mu(t)$ is non-decreasing. To show that $\mu(t)$ is continuous let η be the p-curve defined by $p = p(t)$ for $0 \leqq t \leqq \tau < a$, and let e be an arbitrary positive number. There exists a positive number δ so small that

$$| p(t)p(\tau) | \leqq \frac{e}{2}, \qquad (\tau \leqq t \leqq \tau + \delta \leqq a). \tag{3}$$

Let ζ be the p-curve defined by $p = p(t)$ for $0 \leqq t \leqq \tau + \delta$. Let S_n^2 be an admissible set of n points on ζ. Let S_n^1 be an admissible set of n points on η obtained by replacing each point of S_n^2 for which $t > \tau$ by $p(\tau)$. No point of S_n^2 is thereby moved a distance greater than $e/2$. Hence

$$\begin{aligned} d(S_n^2) &\leqq d(S_n^1) + e, \\ \mu_n(\zeta) &\leqq \mu_n(\eta) + e, \\ \mu(\tau + \delta) &\leqq \mu(\tau) + e. \end{aligned} \tag{4}$$

Since $\mu(t)$ is non-decreasing, (4) implies continuity on the right. Continuity on the left is similarly established, and the proof of (f) is complete.

Statement (g) requires no proof. To establish (h) we assume that there are values τ and τ' of t on $(0, a)$ with $\tau' > \tau$ such that

$p(0)$, $p(\tau)$, and $p(\tau')$ are distinct. Let h and k be the p-curves defined by $p(t)$ for t on $(0,\tau)$ and $(0,\tau')$, respectively. We shall prove that $\mu_k > \mu_h$. To that end, let $2c$ be the minimum value of $p(t)p(\tau)+p(t)p(\tau')$ as t ranges on $(0,\tau)$. We observe that $c \neq 0$ since $p(\tau) \neq p(\tau')$. Let S_n be an admissible set of points $p_1, \cdots, p_n$ on h. We form S'_{n+1} on k by adding one point p_{n+1} to S_n as follows. If

$$p_n p(\tau) \geqq c, \tag{5}$$

we add $p(\tau)$. If (5) does not hold we add $p(\tau')$, and note that

$$p_n p(\tau') \geqq c \tag{6}$$

by virtue of our choice of c. We suppose N so large an integer that $d(S_n) < c$ for $n > N$. For such values of n it follows from (5) and (6) that $d(S'_{n+1}) = d(S_n)$, and hence

$$\mu_{n+1}(k) \geqq \mu_n(h). \tag{7}$$

Since $\mu_n(h)$ is not 0 and tends to 0 as n becomes infinite, for some value of $n > N$,

$$\mu_n(h) > \mu_{n+1}(h), \tag{8}$$

and for such an n it follows from (7) and (8) that $\mu_{n+1}(k) > \mu_{n+1}(h)$. From this relation and from (2) it follows that $\mu_k > \mu_h$ as stated. The proof of (h) is complete.

3. *Equivalent p-Curves.* Let η and ζ be two p-curves given by the respective equations

$$p = p(t), \qquad (0 \leqq t \leqq a), \tag{9}$$

$$q = q(u), \qquad (0 \leqq u \leqq b). \tag{10}$$

Let H be a sense-preserving homeomorphism between the closed intervals $(0, a)$ and $(0, b)$. A homeomorphism of the nature of H will be termed *admissible.* Let $u = u(t)$ be the value of u corresponding to t under H and let $D(H)$ be the maximum of the distances $|p(t)q[u(t)]|$ as t ranges over $(0, a)$. The Fréchet *distance* $\eta\zeta$ between η and ζ is the greatest lower bound of the numbers $D(H)$ as H ranges over the set of all admissible homeomorphisms H between η and ζ. We observe that $\eta\zeta = \zeta\eta \geqq 0$, and one readily proves that for any three p-curves η, ζ, λ, $\eta\lambda \leqq \eta\zeta + \zeta\lambda$.

We shall say that η is *derivable* from ζ if there exists a continuous non-decreasing function $u=u(t)$ which maps the closed interval $(0, a)$ onto the closed interval $(0, b)$ and such that $p(t)\equiv q[u(t)]$. Two p-curves which are derivable from the same p-curve η will be said to be *equivalent*.

Let $\mu(t)$ be the μ-length of η from the point 0 to the point t. Let $t(\mu)$ be the function inverse to $\mu(t)$. To each value of μ on the interval $0\leqq\mu=\mu_\eta$ there corresponds a value or interval of values $t(\mu)$. We set

$$p[t(\mu)] = r(\mu), \qquad (0 \leqq \mu \leqq \mu_\lambda), \tag{11}$$

and observe that $r(\mu)$ is single-valued by virtue of (h), §2. We shall prove the following proposition.

(A) *The function $r(\mu)$ is continuous in μ.*

Let μ_0 be a value of $\mu(t)$. Corresponding to μ_0, let $\tau_1\leqq t\leqq\tau_2$ be the interval of values taken on by $t(\mu)$ at μ_0 (τ_1 may equal τ_2). Corresponding to a positive constant e there exists a positive constant δ so small that

$$\left| p(\tau_2)p(t) \right| < e, \qquad (\tau_2 \leqq t \leqq \tau_2 + \delta). \tag{12}$$

As t ranges over the interval in (12), $\mu(t)$ increases from μ_0 to a value $\mu_1>\mu_0$. From (12) we infer that $\left|r(\mu_0)r(\mu)\right|\leqq e$ for $\mu_0\leqq\mu\leqq\mu_1$. Thus $r(\mu)$ is continuous on the right at μ_0. Continuity on the left is established similarly, and the proof of (A) is complete. We shall now prove the following statement.

(B) *The curve η is derivable from the curve $r=r(\mu)$, $(0\leqq\mu\leqq\mu_\eta)$.*

It follows from the definition of $r(\mu)$ in (11) that $p(t)\equiv r[\mu(t)]$, and the proof of (B) is complete. We term $r=r(\mu)$ a *μ-parameterization* of η and state the following theorem.

THEOREM 1. *Two p-curves are equivalent if and only if they have the same μ-parameterization.*

If η and ζ are equivalent, they are derivable from a common p-curve λ and will have μ-parameterizations identical with that of λ. Conversely, if η and ζ have a common μ-parameterization $r=r(\mu)$, η and ζ are both derivable from $r=r(\mu)$ in accordance with (B) and hence are equivalent. We shall prove the following theorem.

THEOREM 2. *The Fréchet distance between a p-curve ζ and a p-curve η derivable from ζ is null.*

Suppose η and ζ have the representations (9) and (10), respectively. Suppose η is derivable from ζ under the substitution $u=u(t)$, so that $p(t)\equiv q[u(t)]$ for t on $(0, a)$. Let c be an arbitrarily small positive constant and consider the transformation

$$u \equiv [u(t) + ct]\left[\frac{b}{b + ca}\right], \qquad (0 \leqq t \leqq a). \tag{13}$$

This transformation establishes a homeomorphism between the closed intervals $(0, a)$ and $(0, b)$. Denote the right member of (13) by $\phi(t, c)$, and let λ_c be the p-curve

$$q = q[\phi(t, c)], \qquad (0 \leqq t \leqq a). \tag{14}$$

The Fréchet distance $\zeta\lambda_c=0$. For under the transformation (13) corresponding points of ζ and λ_c are identical. To show that $\eta\zeta=0$ we make use of the relation $\eta\zeta \leqq \eta\lambda_c+\lambda_c\zeta=\eta\lambda_c$. If c is sufficiently small, points on η and λ_c determined by the same values of t on $(0, a)$ are arbitrarily and uniformly near since $\phi(t, 0)\equiv u(t)$. Hence $\eta\lambda_c$ tends to 0 with c. But $\eta\zeta$ is independent of c and must then be 0. The proof of the theorem is complete.

We shall now prove Theorem 3.

THEOREM 3. *If η and ζ are two p-curves for which $\eta\zeta<e$, then* $|\mu_\eta-\mu_\zeta| \leqq 2e$.

Let S_n be an admissible set of n points on η. There exists an admissible set S_n' of n points on ζ with distances from the correspondingly numbered points of S_n less than e. Hence $|d(S_n') - d(S_n)| \leqq 2e$, and we infer that $|\mu_n(\zeta) - \mu_n(\eta)| \leqq 2e$. Upon referring to the definition (2) of μ-length we conclude that Theorem 3 holds as stated. We state the following corollary.

COROLLARY 1. *If $\eta\zeta=0$, $\mu_\eta=\mu_\zeta$.*

4. *Curves. A class of equivalent p-curves will be called a curve class or a curve.*

Let α and β be two curves. Let η and η' be p-curves in the class α, and ζ and ζ' p-curves in the class β. I say that

$$\eta\zeta = \eta'\zeta'. \tag{15}$$

Relation (15) follows from the relation

$$\eta\zeta \leqq \eta\eta' + \eta'\zeta' + \zeta'\zeta.$$

For $\eta\eta' = 0$, since η and η' are at a distance 0 from their common μ-parameterized curve λ in accordance with Theorem 2. Similarly $\zeta\zeta' = 0$. Upon reversing the roles of $\eta\zeta$ and $\eta'\zeta'$ we infer that (15) holds as stated. We are accordingly led to the following statement and definition.

The distance between any two p-curve classes α and β equals the distance between any other two p-curves in the classes α and β, respectively, and will be taken as the distance $\alpha\beta$ between the curves α and β.

Let λ be an arbitrary curve with μ-length μ_λ. A pair (λ, μ) will be termed *admissible* if $0 \leqq \mu \leqq \mu_\lambda$. For admissible pairs (λ, μ) let $q(\lambda, \mu)$ be the point on λ which determines the μ-length μ on λ. The following theorem is fundamental.

THEOREM 4. *The point function $q(\lambda, \mu)$ is continuous* in its arguments on the domain of admissible pairs (λ, μ).*

We shall prove (λ, μ) continuous at (λ_0, μ_0) understanding that (λ_0, μ_0) is admissible. Let e be an arbitrary positive constant. We shall show that there exists a positive constant δ such that if (λ, μ) is admissible and

$$\lambda\lambda_0 < \delta, \qquad |\mu - \mu_0| < \delta, \tag{16}$$

then

$$|q(\lambda, \mu)q(\lambda_0, \mu_0)| < e. \tag{17}$$

To that end we shall subject δ to two conditions as follows:

(i) We take $\delta < e/2$. If $\lambda\lambda_0 < \delta$, there will exist a homeomorphism T_δ between μ-parameterizations of λ and λ_0 in which corresponding points have distances less than δ. If the point μ on λ thereby corresponds to μ_1 on λ_0,

$$|q(\lambda, \mu)q(\lambda_0, \mu_1)| < \frac{e}{2}. \tag{18}$$

* We have not yet shown that for two curves α and β, $\alpha\beta = 0$ only if $\alpha = \beta$. But this is not necessary to speak of continuity. The proof of our theorem will imply that $q(\alpha, \mu) = q(\beta, \mu)$ when $\alpha\beta = 0$.

(ii) The second condition on δ is that δ be so small that when $|\mu_1-\mu_0|<3\delta$,

$$|q(\lambda_0, \mu_1)q(\lambda_0, \mu_0)| < \frac{e}{2}. \tag{19}$$

This condition can be satisfied by virtue of the continuity of $q(\lambda_0, \mu)$ in μ.

With δ so chosen I say that (17) holds for admissible pairs (λ, μ) satisfying (16). We introduce T_δ and μ_1 on λ_0 as in (i), and recall that

$$|q(\mu, \lambda)q(\mu_0, \lambda_0)| \leqq |q(\lambda, \mu)q(\lambda_0, \mu_1)| + |q(\lambda_0, \mu_1)q(\lambda_0, \mu_0)|. \tag{20}$$

The first term on the right of (20) is less than $e/2$ by virtue of (18) and the second term is likewise, provided (19) is applicable; that is, provided $|\mu_1-\mu_0|<3\delta$. But under T_δ a point μ on λ will correspond to a point μ_1 on λ_0 such that $|\mu-\mu_1|<2\delta$ in accordance with Theorem 3. Hence $|\mu_1-\mu_0|<3\delta$, (19) is applicable, and the right member of (20) is less than e. The proof of the theorem is complete. We conclude this section with the following theorem.

THEOREM 5. *A necessary and sufficient condition that two p-curves η and ζ have the same μ-parameterization is that $\eta\zeta=0$.*

To prove the condition necessary, let λ represent a μ-parameterized curve determined by η and ζ. Then $\eta\zeta \leqq \eta\lambda+\lambda\zeta$. But $\eta\lambda=\lambda\zeta=0$ by virtue of Theorem 2. Hence the condition is necessary. To prove the condition sufficient suppose that $\eta\zeta=0$ and let $p=p(\mu)$, $(0\leqq\mu\leqq\mu_\eta)$, $q=q(\mu)$, $(0\leqq\mu\leqq\mu_\zeta)$, be μ-parameterizations of η and ζ, respectively. By virtue of Corollary 1 to Theorem 3, $\mu_\eta=\mu_\zeta$. It then follows from Theorem 4 that $p(\mu)$ and $q(\mu)$ differ by a quantity arbitrarily small in absolute value. Hence $p(\mu)=q(\mu)$, and the condition is proved sufficient. Theorem 5 taken with Theorem 1 gives the following corollary.

COROLLARY 2. *A necessary and sufficient condition that two p-curves η and ζ be equivalent is that $\eta\zeta=0$.*

Another way of stating the corollary is to say that for two curves α and β, $\alpha\beta=0$ if and only if $\alpha=\beta$.

INSTITUTE FOR ADVANCED STUDY

A MEAN VALUE THEOREM FOR POLYNOMIALS AND HARMONIC POLYNOMIALS

BY J. L. WALSH

1. *Introduction.* We define a polynomial in z of degree n as any function that can be expressed in the form $a_0z^n+a_1z^{n-1}+\cdots+a_n$; we do not require $a_0\neq 0$. With this definition the following theorems are valid, as is our purpose to show in the present note:

THEOREM 1. *If $f(z)$ is an analytic function of z for the value $z=z_0$, then we have*

$$\lim_{h\to 0}\frac{f(z_0+h)+f(z_0+\omega h)+\cdots+f(z_0+\omega^{N-1}h)-Nf(z_0)}{h^N}=\frac{f^{(N)}(z_0)}{(N-1)!},\tag{1}$$

where ω denotes the number $e^{2\pi i/N}$.

A function $f(z)$ is said to have the *polygonal mean value property* or more simply *the mean value property* if for fixed N and for every z_0 the value of $f(z_0)$ is the mean of the values of $f(z)$ at the vertices of every regular polygon of N sides whose center is z_0.

THEOREM 2. *A necessary and sufficient condition that a function analytic for all values of z have the mean value property is that it be a polynomial of degree $N-1$.*

THEOREM 3. *A necessary and sufficient condition that a real function $f(z)\equiv u(x, y)$ continuous for all values of z $(=x+iy)$ have the mean value property is that it be a harmonic polynomial of degree $N-1$.*

A harmonic polynomial in x and y of degree $N-1$ is defined as the real part of a polynomial in z of degree $N-1$.

2. *Proof of Theorem* 1. In preparation for the proof of Theorem 1 we first formulate a following well known and easily proved lemma.

LEMMA. *If ω denotes the number $e^{2\pi i/N}$, the sum $1+\omega^m+\omega^{2m}+\cdots+\omega^{(N-1)m}$ has the value N or 0 according as m is or is not a multiple of N.*

The result is obvious if m is a multiple of N; in the contrary case we need merely write

$$1+\omega^m+\omega^{2m}+\cdots+\omega^{(N-1)m}=\frac{1-\omega^{Nm}}{1-\omega^m}=0.$$

Let the function $f(z)$ be analytic at the point z_0, hence analytic at every point interior to some circle whose center is z_0. At every such point the Taylor development

$$f(z_0+h)\equiv f(z_0)+\frac{f'(z_0)}{1!}h+\frac{f''(z_0)}{2!}h^2+\cdots$$

is valid, hence by the lemma the relation

$$\begin{aligned}(2)\qquad & f(z_0+h)+f(z_0+\omega h)+\cdots+f(z_0+\omega^{N-1}h)-Nf(z_0)\\ & \equiv N\left[\frac{f^{(N)}(z_0)}{N!}h^N+\frac{f^{(2N)}(z_0)}{(2N)!}h^{2N}+\frac{f^{(3N)}(z_0)}{(3N)!}h^{3N}+\cdots\right]\end{aligned}$$

is also valid. Theorem 1 follows at once from (2), by well known properties of power series.

The special case $N=2$ of Theorem 1 is included in some recent results on real polynomials due to Anghelutza* and Whitney;† the latter writer also gives the corresponding special case of Theorem 2. Both Anghelutza and Whitney deal primarily with the study of *difference* equations, whereas the present properties involve the study of *q-difference* equations.

3. *Proof of Theorem* 2. Every polynomial $f(z)\equiv a_0z^{N-1}+a_1z^{N-2}+\cdots+a_{N-1}$ has the mean value property, for by (2) we may write for every z_0 and h

$$(3)\qquad f(z_0+h)+f(z_0+\omega h)+\cdots+f(z_0+\omega^{N-1}h)\equiv Nf(z_0),$$

which is the mean value property. It is surprising that this simple property of polynomials (also a direct consequence of Lagrange's interpolation formula) is not well known; the writer

* Mathematica, vol. 6 (1932), pp. 1–7.

† This Bulletin, vol. 40 (1934), pp. 89–94.

knows of no explicit statement of it in the literature; the property is contained implicitly in a recent paper by J. H. Curtiss.*

Let $f(z)$ be a function analytic for all values of z, which has the mean value property. The fraction whose limit appears in the left-hand member of (1) is zero for every z_0 and for all values of h different from zero. Consequently we have $f^{(N)}(z) \equiv 0$, and $f(z)$ is a polynomial of degree $N-1$. Theorem 2 is established.

The trivial case $N=1$ is not excluded in Theorem 2; the mean value property requires $f(z_0) \equiv f(z_0+h)$, whence $f(z)$ is a constant.

It is worth remarking that the proof of $f^{(N)}(z) \equiv 0$ just given requires only that (3) should hold for real h; that is to say, it is sufficient if the mean value property holds with reference to all regular polygons with one horizontal radius. Moreover the proof that $f^{(N)}(z) \equiv 0$ requires only that (3) should hold for each $z = z_0$ for h sufficiently small; that is to say, it is sufficient if the mean value property holds for each z_0 with reference to regular polygons which are sufficiently small. Consequently if $f(z)$ is analytic in a region and has the mean value property with reference to all polygons contained in that region, then in that region $f(z)$ is a polynomial of degree $N-1$.

If a function is analytic in a region and possesses the mean value property for given N for all sufficiently small polgyons of N sides in that region, the function also possesses the mean value property for all sufficiently small polygons of $N+1$, $N+2, \cdots$ sides in that region.

Let a function $f(z)$ be analytic in the neighborhood of a *fixed* point z_0 and possess the mean value property merely with reference to all sufficiently small regular polygons of N sides *with center* z_0. It follows then from (3) and (2) that we have $0 = f^{(N)}(z_0) = f^{(2N)}(z_0) = f^{(3N)}(z_0) = \cdots$. Reciprocally, if $f(z)$ is analytic in a circle whose center is a fixed point z_0 and if we have $0 = f^{(N)}(z_0) = f^{(2N)}(z_0) = f^{(3N)}(z_0) = \cdots$, then it follows from (2) and (3) that $f(z)$ has the mean value property with reference to all regular polygons of N sides with center z_0 which lie in the given circle.

4. *Proof of Theorem* 3. Let $u(x, y)$ be a harmonic polynomial of degree $N-1$; the conjugate function $v(x, y)$ is also a har-

* Transactions of this Society, vol. 38 (1935), pp. 458–473; p. 462.

monic polynomial of degree $N-1$, and the analytic function $f(z) \equiv u(x, y)+iv(x, y)$ is a polynomial in z of degree $N-1$. The function $f(z)$ has the mean value property (3). If we take the real part of both members of (3) we obtain the mean value property for $u(x, y)$.

An arbitrary function $U(x, y)$ is said to have the *Gauss mean value property* if for every (x_0, y_0) the value $U(x_0, y_0)$ is the mean of the values of $U(x, y)$ over an arbitrary circumference whose center is (x_0, y_0). Every function harmonic in a region possesses the Gauss mean value property with respect to circumferences which together with their interiors lie in that region. Conversely [Koebe*], if a function $U(x, y)$ is continuous in a region and has the Gauss mean value property with respect to all sufficiently small circumferences in that region, then $U(x, y)$ is harmonic in that region.

Let now $u(x, y)$ be continuous in a region or throughout the entire plane, and possess the polygonal mean value property at least for sufficiently small polygons. Allow such a sufficiently small polygon to rotate rigidly about its center through an angle $2\pi/N$; the average of $u(x, y)$ over the entire circumscribed circle is the average of $u(x, y)$ at the N vertices averaged during this rotation, and is therefore equal to the value of $u(x, y)$ at the center of the circle. That is to say, the function $u(x, y)$ possesses the Gauss mean property, at least with respect to sufficiently small circumferences, and hence is harmonic.

Let $v(x, y)$ be conjugate to $u(x, y)$ (continuous and possessing the polygonal mean value property) in the given region or in a simply-connected subregion. Then $f(z)=u(x, y)+iv(x, y)$ is analytic there. Let us use equation (1) where h is chosen as real. At an arbitrary point z we have

$$f^{(N)}(z) = \frac{\partial^N u}{\partial x^N} + i\,\frac{\partial^N v}{\partial x^N}.$$

From the polygonal mean value property for $u(x, y)$ it follows by Theorem 1 with h real that $\partial^N u/\partial x^N$ (the real part of $f^{(N)}(z)$) vanishes identically, so the function $\partial^N v/\partial x^N$ conjugate to $\partial^N u/\partial x^N$ either vanishes identically or is identically a non-vanishing real constant. If $\partial^N v/\partial x^N$ vanishes identically, so also

* See for instance Kellogg, *Potential Theory*, 1929, pp. 224–228.

does $f^{(N)}(z)$, and Theorem 3 is established. If $\partial^N v/\partial x^N$ is identically a non-vanishing real constant q, we have $f^{(N)}(z) \equiv iq$,

$$f(z) \equiv u + iv \equiv \frac{iqz^N}{N!} + a_1 z^{N-1} + a_2 z^{N-2} + \cdots + a_N;$$

we shall reach a contradiction. Let $z = z_0$ be an arbitrary point of the plane or more generally of the region in which $u(x, y)$ is assumed to have the polygonal mean value property. We write

$$f(z) \equiv \frac{iq(z - z_0)^N}{N!} + b_1(z - z_0)^{N-1} + b_2(z - z_0)^{N-2} + \cdots + b^N.$$

The real part of $b_1(z-z_0)^{N-1}+b_2(z-z_0)^{N-2}+ \cdots +b_N$ is known to have the polygonal mean value property, hence the real part of $F(z) \equiv iq(z-z_0)^N/N!$ also has the polygonal mean value property; this is not true as we see by setting $z = z_0+\alpha h$, where h is real and positive and where α is an Nth root of $-i$; we have

$$F(z_0 + \alpha h) + F(z_0 + \alpha\omega h) + F(z_0 + \alpha\omega^2 h) + \cdots + F(z_0 + \alpha\omega^{N-1} h) - NF(z_0) = qh^N/(N - 1)!,$$

whose real part is not zero. Thus $q \neq 0$ leads to a contradiction and Theorem 3 is completely established. We have indeed proved more than the italicized theorem, for we have shown that the polygonal mean value property at each point of a region for sufficiently small polygons implies that $u(x, y)$ is in that region a harmonic polynomial of degree $N-1$.

5. *Extensions of Theorem* 3. We have seen that the polygonal mean value property for sufficiently small polygons for a *continuous* function $u(x, y)$ implies that $u(x, y)$ is a harmonic polynomial of degree $N-1$. It is essential here to assume $u(x, y)$ continuous, as we see by the example

$$u(x, y) = \begin{cases} 1, & x > 0, \\ 0, & x = 0, \\ -1, & x < 0, \end{cases}$$

where N is even. At each point (x, y) the polygonal mean value property is possessed by this function for sufficiently small polygons. Yet the function is not a harmonic polynomial.

The continuity of $u(x, y)$ need not be assumed, however, if the polygons are not required to be small. More explicitly, let us indicate that *if* $u(x, y)$ (*not assumed continuous but measurable and Lebesgue integrable both on every circumference and interior to every circle in* R) *has the mean value property at all points of a region* R *with reference to all* N*-sided polygons in that region, then* $u(x, y)$ *is in that region a harmonic polynomial of degree* $N-1$. The polygonal mean value property yields the Gauss mean value property, as we have already seen. The Gauss mean value property implies by integration over the interior of any circle interior to R that the average of $u(x, y)$ over the interior of any such circle is the value of $u(x, y)$ at the center of the circle. The latter property implies continuity of $u(x, y)$. For let $P_0: (x_0, y_0)$ be an arbitrary point interior to R and let D in R be the interior of a circle whose center is (x_0, y_0) and radius r_0. Let $P_k: (x_k, y_k)$ be an arbitrary sequence of points approaching (x_0, y_0), and let D_k be (for k sufficiently large) the interior of a circle whose center is (x_k, y_k) and radius $r_0 - \overline{P_0P_k}$. Then D_k lies interior to D; the radius of D_k approaches the radius of D; by Lebesgue's theorem the average value of $u(x, y)$ over D_k approaches the average value of $u(x, y)$ over D. Consequently $u(x, y)$ is continuous at the point (x_0, y_0). Theorem 3 extended now implies that $u(x, y)$ is in R a harmonic polynomial of degree $N-1$.

6. *Geometric Mean Values.* It is not without interest to compare our fundamental q-difference equation (3) with the analogous equation

$$(4) \qquad \phi(z_0 + h)\phi(z_0 + \omega h) \cdots \phi(z_0 + \omega^{N-1}h) \equiv [\phi(z_0)]^N,$$

which expresses the condition that the value of the function $\phi(z)$ at a point z_0 is the geometric mean of the values at the vertices of a regular N-sided polygon whose center is z_0. Equation (4) reduces to (3) by taking the logarithms of both members of (4), so we have by Theorem 2 and its extension:

THEOREM 4. *A necessary and sufficient condition that an analytic function* $\phi(z)$ *have at every point* z_0 *of a region the mean value property* (4) *for all values of* h *or for all* h *sufficiently small is that* $\phi(z)$ *be identically zero or be of the form* $e^{p(z)}$, *where* $p(z)$ *is a polynomial of degree* $N-1$.

It will be noticed that it is possible to take logarithms in (4) in the neighborhood of a point z_0 if $\phi(z)$ is analytic, is not identically zero, and satisfies (4); if (4) is satisfied for every h sufficiently small we cannot have $\phi(z_0)=0$ unless at least one of the factors in the left-hand member of (4) vanishes for every h sufficiently small; in the latter case z_0 is a non-isolated zero of $\phi(z)$ and $\phi(z)$ vanishes identically.

Let now the real continuous function $U(x, y)$ have the property that its value at every point (x_0, y_0) is the geometric mean of its values at the vertices of every N-sided regular polygon whose center is (x_0, y_0). If $U(x, y)$ vanishes at a single point (x_1, y_1), it vanishes at every point, for we may choose a regular polygon of N sides whose center is (x_0, y_0) and which has a vertex at (x_1, y_1). Even if the mean value property is restricted to all polygons which are sufficiently small, an extension of the method of proof of Theorem 4 will yield the corresponding result:

THEOREM 5. *A necessary and sufficient condition that a real continuous function $U(x, y)$ have in a region or in the entire plane the property that the value at the center of every sufficiently small regular polygon of N sides is the geometric mean of the values at the vertices, is that $U(x, y)$ be identically zero or be of the form $\pm e^{u(x,y)}$ (the lower sign is permissible only if N is odd), where $u(x, y)$ is a harmonic polynomial of degree $N-1$.*

The trivial case $U(x, y)\equiv 0$ is henceforth excluded. Let now R be the given region, and let R' be one of the subregions of R obtained by deleting from R the points where $U(x, y)$ vanishes. The function $\log U(x, y)$ or $\log[-U(x, y)]$ (according as $U(x, y)$ is positive or negative in R') is continuous in R' and has the mean value property expressed by (3), so $U(x, y)$ is of the form $\pm e^{u(x,y)}$, where $u(x, y)$ is a harmonic polynomial of degree $N-1$. The function $u(x, y)$ cannot become negatively infinite at any finite boundary point of R', so $U(x, y)$ does not vanish at any finite boundary point of R', and R' may be chosen identical with R. Theorem 5 is established.

If the requirement of continuity is omitted, and if for each point only sufficiently small polygons are allowed, the conclusion is false, as we see by the illustration (N even)

$$U(x, y) = \begin{cases} e, & x > 0, \\ 1, & x = 0, \\ e^{-1}, & x < 0. \end{cases}$$

The analogs of Theorems 3 and 5 in three or more dimensions require methods other than those developed in the present note, and are postponed for another occasion.

HARVARD UNIVERSITY

ON SEQUENCES OF INDEFINITE INTEGRALS

BY M. K. GOWURIN

1. *Introduction.* The chief result concerning the subject of this paper is due to Lebesgue* and can be formulated as follows:

If $\{f_n(t)\}$ is a sequence of functions defined and integrable in $J=(0, 1)$ and if for every measurable set $e \subset J$

$$\lim_n \int_e f_n(t)dt = 0, \tag{1}$$

then the sequence of indefinite integrals

$$\int_e f_n(t)dt \tag{2}$$

is uniformly absolutely continuous.

G. Fichtenholz† has shown that the same conclusion remains true, if the equality (1) is satisfied for all open sets e.

S. Saks‡ considered the space $R=\{x\}$ of the characteristic

---

* *Sur les intégrales singulières*, Annales de Toulouse, (3), vol. 1 (1909), p. 58; see also Hahn, *Über Folgen linearen Operationen*, Monatshefte für Mathematik und Physik, vol. 32 (1922), p. 45.

† *Theory of simple definite integrals depending on a parameter*, Petrograd, 1918, p. 98 (in Russian) or *Sur les suites convergentes des intégrales définies*, Bulletin de l'Académie des Sciences de Pologne, Sér. A, Décembre, 1923, pp. 115, 117.

‡ *On some functionals*, Transactions of this Society, vol. 35 (1933), pp. 549–556.

functions of measurable sets contained* in J, by introducing the metric:

$$\|x\| = \int_0^1 x(t)dt = mE_t(x(t) = 1), \quad d(x, y) = \int_0^1 |x(t) - y(t)|\, dt.$$

He proved the following generalization of Lebesgue's theorem:

If the equality (1) *written in the form*

$$\lim_n F_n(x) = \lim_n \int_0^1 f_n(t)x(t)dt = 0, \tag{1'}$$

is satisfied for all points x belonging to a certain set H of the second category in R, then the sequence of indefinite integrals (2) *is uniformly absolutely continuous, that is, the sequence of functionals*

$$F_n(x) = \int_0^1 f_n(t)x(t)dt \tag{2'}$$

is equally continuous in R, and $F_n(x) \to 0$ for all $x \in R$.

G. Fichtenholz has stated the following problem: *What is the general characteristic of the sets H in R, such that the equal continuity of the sequence* (2') *follows from its convergence to zero*† *on H?* I shall study in particular the role of the category of H in this problem.‡

---

* The space R is complete but not linear. The sum of two elements of R, $x+y$, exists if $x(t)\cdot y(t)=0$, their difference $x-y$ exists if $x(t)\cdot y(t)=y(t)$. (Here as everywhere, all the equalities referring to the characteristic functions must be satisfied almost everywhere.) Further on we shall not prove the existence of sums and differences occurring in the text, since it is evident from the course of reasoning. The *zero* in the space R is the characteristic function of the vacuous set. Note also the equality $\|x+y\| = \|x\| + \|y\|$. Obviously, J can denote not only an interval, but every measurable set. We shall apply this statement in §3.

† As he showed, the convergence to zero of the sequence (2') on all the R always follows from its convergence to zero on such a set H.

‡ [Added in proof.] As Saks has kindly pointed out to me, the totality of all the open sets considered by Fichtenholz is of the first category in R. Nevertheless, it seems to me that the example stated below is not without interest.

Without solving the problem stated above, I give in this paper an example which shows that in any case the category H is not the decisive factor in the question under consideration. Of course it is not at all difficult to construct an example, of a set H which is of the first category in R, but, being enlarged by all the existing linear combinations of its elements, forms a set of the second category in R. It is more natural, however, to consider the set H as being additive (that is, such that every existing linear combination of its elements belongs to it again). We shall prove the following statement.

THEOREM. *If the equality* (1′) *is satisfied for all points x belonging to some spherical surface in R, $S^1(x_0, r)$, $(r<1)$, which does not pass through the point* 0, *then the system* (2′) *is equally continuous in R and $F_n(x)\to 0$ everywhere in R.*

In general, the set $H=S(x_0, r)$ is not additive, but it is easy to see that the additive extension of $S(x_0, r)$ is nowhere dense in R.

2. *Proof of Theorem.* Let $x_0=0$. Then the spherical surface $S(0, r)$ consists of the points of norm r. Let us suppose that our assertion is wrong and that there exists a system (2′) converging to zero on $S(0, r)$, but not equally continuous. There exists an ϵ such that, for every $\delta>0$ and every integer N, R will contain an element x such that $\|x\|<\delta$ and $|F_n(x)|>\epsilon$ for a certain $n\geqq N$. Denote by H_m the set of points x of $S(0, r)$ such that $|F_n(x)|\leqq\epsilon/6$ for $n>m$. Evidently all H_m are closed and as $\sum_1^\infty H_m=S(0, r)$, at least one of them, for instance H_{m_0}, is of the second category in $S(0, r)$. Being closed, H_{m_0} contains in $S(0, r)$ a sphere $T(0, r; \xi_0, \rho)$ or, more briefly, T (T consists of such points x that $\|x\|=r$ and $d(x, \xi_0)\leqq\rho$).

According to the assumption there exists in R an element x_1 such that

$$\|x_1\| = \alpha_1 < \frac{r}{3}, \qquad \alpha_1 < \frac{1-r}{3}, \qquad \alpha_1 < \frac{\rho}{4},$$

and for some $\nu>m_0$, $|F_\nu(x_1)|>\epsilon$. For example, let $F_\nu(x_1)>\epsilon$.

Consider an arbitrary element $\bar{x}$ of the spherical surface $S(0, \alpha_1)$. Construct such an element z that $z+x_1$ and $z+\bar{x}$ exist and are contained in T. The method of construction de-

pends upon the quantity $\beta = \|(x_1+\bar{x}-x_1\bar{x})\xi_0\|$, which can be equal to, greater, or less than α_1.

In the first case $z=\xi$, where $\xi=\xi_0-(x_1+\bar{x}-x_1\bar{x})\xi_0$.

In the second case we put $z=\xi+\eta$, where η is the characteristic function of a set of measure $\beta-\alpha_1$, lying in $E_t(\xi_0(t)+x_1(t)+\bar{x}(t)=0)$. It follows from $\beta-\alpha_1 \leqq \alpha_1$ and from

$$mE_t(\xi_0(t)+x_1(t)+\bar{x}(t)=0) \geqq 1-(\|\xi_0\|+\|x_1\|+\|\bar{x}\|)$$
$$= 1-r-2\alpha_1 \geqq 1-r-2\frac{1-r}{3} = \frac{1-r}{3} > \alpha_1,$$

that such a set exists.

In the third case $z=\xi-\chi$, where χ is the characteristic function of a set of measure $\alpha_1-\beta$ lying in $E_t(\xi(t)=1)$. It follows from

$$mE_t(\xi(t)=1) = \|\xi\| \geqq r-2\alpha_1 > \alpha_1,$$

that such a set exists.

In each of these cases $\|z\| = r-\alpha_1$ and $d(z, \xi_0) \leqq 3\alpha_1$. Evidently z satisfies the conditions indicated above. Then

$$F_\nu(z+x_1) = F_\nu(z)+F_\nu(x_1), \qquad F_\nu(z+\bar{x}) = F_\nu(z)+F_\nu(\bar{x});$$
$$F_\nu(\bar{x}) = F_\nu(z+\bar{x})-F_\nu(z) = F_\nu(z+\bar{x})-F_\nu(z+x_1)+F_\nu(x_1).$$

As $z+x \in T$ and $z+\bar{x} \in T$, the first two members of the right part of the last equality are numerically $<\epsilon/6$, whence $F_\nu(\bar{x})>4\epsilon/6$.

Note that $\bar{x}$ is an arbitrary element of $S(0, \alpha_1)$. Let α_2 be the least non-negative residue of r mod α_1, that is, $r=k_1\alpha_1+\alpha_2$, (k_1 a positive integer, $0 \leqq \alpha_2 < \alpha_1$). Then ξ_0 can be represented in the form

$$\xi_0 = u_1+u_2+\cdots+u_{k_1}+x_2,$$

where $u_i \cdot u_j = 0$, $(i \neq j)$, $u_i x_2 = 0$, $\|u_i\| = \alpha_1$, and $\|x_2\| = \alpha_2$.

We have

$$F_\nu(x_2) = F_\nu(\xi_0)-F_\nu(u_1)-\cdots-F_\nu(u_{k_1}).$$

Since $|F_\nu(\xi_0)| < \epsilon/6$ and $F_\nu(u_i) \geqq 4\epsilon/6$, we have

$$F_\nu(x_2) < \frac{\epsilon}{6}-k_1\frac{4\epsilon}{6} \leqq -\frac{3\epsilon}{6}.$$

Repeating for x_2 and $S(0, \alpha_2)$ the reasoning given for x_1 and $S(0, \alpha_1)$, we find for $x \in S(0, \alpha_2)$, $F_\nu(x) < -\epsilon/6$.

Dividing α_1 by α_2, we obtain

$$\alpha_1 = k_2\alpha_2 + \alpha_3, \qquad (k_2 \text{ a non-negative integer}, 0 \leqq \alpha_3 < \alpha_2).$$

If $x \in S(0, \alpha_3)$, then adding to it k_2 suitably selected elements $v_1, v_2, \cdots, v_{k_2}$ belonging to $S(0, \alpha_2)$, $(v_i \cdot v_j = 0$, when $i \neq j$, and $v_i \cdot x = 0)$, we obtain an element $y \in S(0, \alpha_1)$. Then

$$F_\nu(x) = F_\nu(y) - F_\nu(v_1) - \cdots - F_\nu(v_{k_2}),$$

and

$$F_\nu(x) > \frac{4\epsilon}{6} + k_2 \frac{\epsilon}{6} \geqq \frac{5\epsilon}{6}.$$

Divide α_2 by α_3, and so on. If r and α_1 are incommensurable, we shall obtain a sequence of spheres $S(0, \alpha_p)$ converging to zero, where

$$\begin{aligned}
&\text{on } S(0, \alpha_1), && F_\nu(x) > 4\epsilon/6,\\
&\text{on } S(0, \alpha_2), && F_\nu(x) < -\epsilon/6,\\
&\text{on } S(0, \alpha_3), && F_\nu(x) > 5\epsilon/6,\\
&\text{on } S(0, \alpha_4), && F_\nu(x) < -6\epsilon/6,\\
&\text{on } S(0, \alpha_5), && F_\nu(x) > 11\epsilon/6,
\end{aligned}$$

and so on. Such a result contradicts the continuity of $F_\nu(x)$. If r and α_1 are commensurable, a certain α_μ will be a divisor of r. Then

$$\xi_0 = w_1 + w_2 + \cdots + w_\lambda,$$

where $w_i \cdot w_j = 0$, $(i \neq j)$, and $w_i \in S(0, \alpha_\mu)$, $(i = 1, 2, \cdots, \lambda)$. The numbers $F_\nu(w_i)$ have the same sign for every i and $|F_\nu(w_i)| > \epsilon/6$. Consequently, $|F_\nu(\xi_0)| > \epsilon/6$, which is impossible, since $\nu > m_0$ and $\xi_0 \in H_{m_0}$. Thus the system (2′) is really equally continuous.

Thus every sequence (2′), which satisfies condition (1′) on $S(0, r)$ is equally continuous. Hence it follows that such a sequence converges to zero everywhere* on R, according to the general assertion.

* The reader sees that as far as it concerns the spherical surface $S(0, r)$ with the center at the point $x = 0$, the proof will not have to be changed considerably if we suppose that H, although not coinciding with $S(0, r)$, is of the second category in $S(0, r)$.

3. *Application of Functionals.* Now let $S(x_0, r)$ be any spherical surface which does not pass through 0 and suppose that, for $x \in S(x_0, r)$, $\lim_n F_n(x) = 0$. Note that every spherical surface in R has two centers and divides the whole space into two spheres, that is, that $S(x_0, r) = S(1-x_0, 1-r)$. Suppose that x_0 is that one of the two centers whose norm is greater than the radius.

For further reasoning it is necessary to show that $F_n(x_0) \to 0$. For that purpose consider the set of points $x \in S(x_0, r)$ for which $x \cdot x_0 = x$ and denote it by S'. If we introduce the set $J' = E_t(x_0(t) = 1)$, then $S' = S'(0, \|x\| - r)$ is a spherical surface with the center at 0 and with radius $\|x_0\| - r$ in the space R' of characteristic functions of measurable subsets of J'. Since reasoning in §2 remains valid when J' is substituted for J, we have $F_n(x) \to 0$ everywhere in R' and in particular at the point x_0.

Any element $x \in R$ can be represented in the form $x = x_0 - u(x) + v(x)$, where $u(x) = x_0(1-x)$ and $v(x) = x(1-x_0)$. Let us define a one-to-one and bicontinuous transformation of R into itself according to the formulas

$$y = \Psi(x) = u(x) + v(x) = x_0(1 - x) + x(1 - x_0),$$

$$x = \Psi^{-1}(y) = x_0 - yx_0 + y(1 - x_0).$$

In this way $S(x_0, r)$ is transformed into $S(0, r)$.

Further define in the transformed space the sequence of functionals $\{\Phi_n(y)\}$ on putting

$$\text{(A)} \qquad \Phi_n(y) = F_n(x) - \frac{F_n(x_0)}{\|x_0\|} \int_0^1 x_0(t)x(t)dt.$$

If we remember the definition of $F_n(x)$ and express x in terms of y, we easily find that the functionals $\Phi_n(y)$ have the form required,

$$\Phi_n(y) = \int_0^1 g_n(t)y(t)dt.$$

Since $F_n(x_0) \to 0$ and $F_n(x) \to 0$ on $S(x_0, r)$, we conclude from the equality (A) that $\Phi_n(y)$ converges to zero on $S(0, r)$. Applying the result of §2 to the sequence $\{\Phi_n(y)\}$, we see that

$\Phi_n(y) \to 0$ everywhere in R. Since the transformation is one-to-one, it follows from the formula (A) that $F_n(x) \to 0$ everywhere in R.

Note that the statement made in §1 is not true for a spherical surface which passes through zero, $S(x_0, \|x_0\|)$. This we can prove on putting

$$F_n(x) = n \int_0^1 [x_0(t) - (1 - x_0(t))]x(t)dt.$$

Let $x \in S(x_0, \|x_0\|)$; $x = x_0 - u + v$ and

$$\begin{aligned} \|x_0\| = d(x, x_0) &= \int_0^1 | x(t) - x_0(t) | dt \\ &= \int_0^1 | - u(t) + v(t) | dt \\ &= \int_0^1 [u(t) + v(t)]dt = \|u\| + \|v\|, \end{aligned}$$

that is, $\|x_0 - u\| = \|v\|$. According to definition, $x_0 - u = x_0 \cdot x$ and $v = x(1 - x_0)$. Consequently for such an x, $F_n(x) = 0$, for $n = 1, 2, \cdots$. At the same time the sequence $\{F_n(x)\}$ is by no means equally continuous in R, as $|F_n(x)| \to \infty$ for every x non-$\in S(x_0, \|x_0\|)$.

4. *Remark.* In the paper of Saks mentioned above there are some vague points in the proof of Lemma 4. This refers to the way he motivates the inclusion (on page 553)

$$H \subset H_1 + H_2 + \cdots.$$

The lemma itself is not true, at least for the space R, as is shown by the example of the following sequence of transformations:

$$\xi_n(x, t) = nx(t).$$

Moreover, the assertions of Theorems 3 and 4 are not true in so far as they refer to R, as is seen from the same example.

INSTITUTE OF MATHEMATICS,
LENINGRAD UNIVERSITY

THE WEB OF QUADRICS

BY T. R. HOLLCROFT

1. *Introduction.* A triply infinite linear system of quadrics was first defined by De Jonquières* in 1862. Such a linear system later became known as a web.

The characteristics of a web of quadrics, with and without basis elements, have never been completely or accurately given, although many mathematicians have investigated this problem during the past seventy-five years.† Also, the web with a basis double point, that is, the web of quadric cones, has not been treated.

The purpose of this paper is to give a brief summary of results, correcting errors and completing the sets of characteristics; also, to derive the characteristics of a web of quadric cones. References to more complete discussions of certain parts are given.

In all cases, the quadrics will be considered as loci of points. Each web type has its dual, whose treatment in plane coordinates is algebraically identical with that in point coordinates.

2. *The Web and the Associated Involution.* The web of quadrics has the equation

$$\sum \lambda_i f_i = 0, \qquad (i = 1, 2, 3, 4),$$

in which the $f_i = 0$ are equations of quadrics with given basis elements. The web has a jacobian quartic surface J.

The quadrics of the web in a 3-space (x) are in (1, 1) correspondence with the planes of a 3-space (y). The correspondence is defined by the equations

$$\rho y_i = f_i, \qquad (i = 1, 2, 3, 4).$$

---

* De Jonquières, Journal de Mathématiques, (2), vol. 7 (1862), p. 412.

† Encyklopädie der Mathematischen Wissenschaften, vol. III_2, pp. 250–254. Pascal, *Repertorium der höheren Mathematik*, vol. II_2 (1922), pp. 629–631. Virgil Snyder and F. R. Sharpe, *Space involutions defined by a web of quadrics*, Transactions of this Society, vol. 19 (1918), pp. 275–290.

This correspondence gives rise to an involution between the points of (y) and (x). The order of this involution is the number of non-basic intersections of three quadrics of the web. In the involution, the branch-point surface L in (y) is in (1, 1) correspondence with J, which is the coincidence surface in (x). This involution is fully discussed in the paper by Snyder and Sharpe cited above.

The complete image of J is L. The complete image of L is J counted twice and a residual surface R. The properties of the web are defined by the characteristics of the three surfaces L, J, and R.

The web contains ∞^3 nets and ∞^4 pencils of quadrics. The systems are in (1, 1) correspondence with the points (vertices of bundles of planes) and lines (axes of pencils of planes) of (y), respectively. The jacobian curves of the nets of quadrics form a web of curves on J, whose images on L are the contour curves of the associated tangent cones to L. Two jacobian curves of any two nets of the web have intersections whose number is equal to the order of the discriminant of a quadric of the web.

The characteristics of $L, J,$ and R are obtained by the methods of a former paper.* It is to be noted that, since the quadric web contains composite surfaces and this does not occur for $n>2$, the quadric web is a special web whose characteristics do not all result on substituting $n=2$ in the formulas for a general n.

For a web of quadrics, L has neither bitangent planes nor stationary tangent planes and therefore neither a bitangential nor a parabolic curve.

The characteristics of L that do not vanish are represented by the following symbols: N, order; a, order of tangent cone; a', class of plane section; n', class; b, order of nodal curve; c, order of cuspidal curve; ρ, class of nodal curve; σ, class of cuspidal curve; γ, number of intersections of b and c which are cusps on b;† β, number of intersections of b and c which are cusps on c; q, rank of b; r, rank of c; p_1, genus of plane section; δ', bitangents of plane section; κ', stationary tangents of plane section; δ, nodal generators of tangent cone; κ, cuspidal generators of tan-

* T. R. Hollcroft, *The general web of algebraic surfaces of order n and the involution defined by it*, Transactions of this Society, vol. 35 (1933), pp. 855–868.

† The same symbol will be used both as the order and name of a curve.

gent cone; C, number of conic nodes; C', number of conic tropes.

The image of b is the intersection curve b_1 (counted twice) of J and R and the nodal curve b_2 of R. The image of c is the contact curve c_1 (counted three times) of J and R and a residual curve c_2 which is the cuspidal curve of R.

3. *The Web with Simple Basis Points.* The web of quadrics may have any number up to six of independent simple basis points. For any number of basis points, zero to six inclusive, the characteristics of L are given in the following table.

No. basis pts. / Char. of L	0	1	2	3	4	5	6
N	16	14	12	10	8	6	4
n'	4	4	4	4	4	4	4
$a=a'$	12	12	12	12	12	12	12
p_1	9	8	7	6	5	4	3
δ'	22	23	24	25	26	27	28
κ'	24	24	24	24	24	24	24
b	60	40	24	12	4	0	0
c	36	30	24	18	12	6	0
γ	120	72	36	12	0	0	0
β	80	56	36	20	8	0	0
t	80	32	8	0	0	0	0
q	40	28	18	10	4	0	0
r	68	62	54	44	32	18	0
ρ	80	56	36	20	8	0	0
σ	32	32	30	26	20	12	0
δ	28	28	28	28	28	28	28
κ	24	24	24	24	24	24	24
C	0	0	1	3	6	10	16
C'	10	11	12	13	14	15	16

The corresponding characteristics of J and R are given in the following table. In this table, "mult. P_i on J" means "the multiplicity of *each* simple basis point on J," and similarly for R and the curves.

When the web has six basis points, J is a Weddle quartic surface and L is a Kummer quartic surface.*

No. basis points / $J, R,$ in (x).	0	1	2	3	4	5	6
order of J	4	4	4	4	4	4	4
mult. P_i on J		2	2	2	2	2	2
no. lines, J	10	10	11	13	16	20	25
order of R	24	20	16	12	8	4	0
mult. P_i on R		10	8	6	4	2	0
order of c_1	18	18	17	15	12	8	0
mult. P_i on c_1		6	5	4	3	2	0
order of b_1	60	44	30	18	8	0	0
mult. P_i on b_1		8	6	4	2	0	0
order of c_2	90	66	45	27	12	0	0
mult. P_i on c_2		12	9	6	3	0	0
order of b_2	120	72	36	12	0	0	0
mult. P_i on b_2		24	12	4	0	0	0

The image of each basis point is a plane of (y) whose image in (x) is the basis point and the quadric tangent cone to J at the basis point. All lines joining basis points are lines of J.

4. *The Web with Basis Curves.* The web of quadrics may have the following basis curves or combinations of simple basis points and basis curves:

(1) *Web with basis line l.* The surface J is a ruled quartic surface with the line l three-fold. Every line of J is the axis of a composite quadric. The only possible cones of the web are those whose vertices lie on l. The surface L is the developable of the curve c, and has the characteristics: $N=6$, $a=a'=4$, $n'=0$, $b=4$, $c=6$, $\kappa'=0$, $\delta'=3$, $\kappa=0$, $\delta=0$, $\sigma=4$, $\rho=8$, $\beta=4$, $\gamma=0$, $t=0$, $r=6$, $q=6$. R is a ruled quartic that does not contain l. The curve c_1 is of order 5 and intersects l in four points. The curve b_1 is of order 6 and intersects l in the same four points, which are the

* For a derivation of the equations of J and L and further particulars of this case, see Hudson, *Kummer's Quartic Surface*, 1905, pp. 169–172.

intersections of l and R. The cuspidal curve of R is of order 3. R has no nodal curve. The image of l is a quadric of (y) whose generators correspond to the points of l.

(2) *Web with two basis skew lines l_1 and l_2.* This web contains no cones and any two planes through l_1 and l_2, respectively, constitute a composite quadric of the web. J is, therefore, indeterminate. The lines l_1 and l_2 belong to the same regulus for each proper quadric. Three quadrics of the web have no intersections outside of the two basis lines and thus no involution is established.

(3) *Two coplanar lines l_1 and l_2, intersecting at P.* For all quadrics of the web, l_1 and l_2 belong to different reguli and the plane of these two lines is a common tangent plane at P.

The only cones of the web have their vertices at P and contain l_1 and l_2 as generators. The axes of composite quadrics generate a quadric cone with vertex at P and containing l_1 and l_2. This quadric cone and the plane of the basis lines counted twice constitute the surface J. The involution is of order two and L is a quadric surface.

(4) *A basis conic c.* J consists of the plane π of c counted twice and a quadric surface through c. L is a quadric. The involution is of order two. The web contains no composite quadrics except π^2. Each cone of the web has its vertex and two generators on the quadric which is the proper J.

(5) *One line and one or two basis points.* If the web has one basis line l and one basis point P, J consists of a ruled cubic surface through P with l double and the plane determined by l and P. The proper J is the cubic surface. L is a ruled quartic surface with a cuspidal cubic and no nodal curve. The involution is of order three. R is a quadric cone with vertex at P; and J and R have a contact cubic c_1 through P.

If the web has one basis line l and two basis points P_1 and P_2, J consists of a quadric through l and both basis points and two planes, both through l and one each through P_1 and P_2. The involution is of order two. L is a quadric.

If there are additional basis elements, such as seven points, one line and three points, two coplanar lines or a conic and one point, the web is homaloidal and the associated involution is birational. The characteristics of such webs are well known.

5. *Properties of the Web.* The following discussion applies to all the preceding cases.

J is the locus of vertices of cones, of axes of composite quadrics, and of all contacts of quadrics of the web.

The contact curve c_1 of J and R is the locus both of three-point coincidences of the involution and of stationary contacts of pencils of quadrics of the web. The curves c and c_1 are in $(1, 1)$ correspondence.

The intersection curve b_1 of J and R is the locus both of pairs of coincidences of the involution and of pairs of simple contacts of pencils of quadrics of the web. The curves b and b_1 are in $(1, 2)$ correspondence.

The points γ correspond to pairs of points P_1P_2 on b_1, through one of which, P_1, passes c_1. In the involution, to a point γ corresponds a three-point coincidence at P_1 and a simple coincidence at P_2. In the web, there are γ pencils of quadrics in which the quadrics have stationary contact at a point P_1 and simple contact at the associated point P_2.

The points β correspond to points of osculation of b_1 and c_1. Such a point is a four-point coincidence in the involution. In the web, there are β pencils of quadrics in which the quadrics have tacnodal contact at a point of osculation of b_1 and c_1.

The points t correspond to t triplets of nodes on b_1. Each triplet is a set of three simple coincidences in the involution. In the web, there are t pencils of quadrics such that in each pencil the quadrics have three simple contacts at the three points corresponding to a triple point of b.

In case the surface L does not have certain singularities, the quadrics of the associated web do not have the contacts to which these singularities correspond. Further details are given in the paper cited in §2.

6. *The Web with a Double Basis Point.* The web of quadrics with a double basis point Q is a web of quadric cones with a common vertex, at Q. This web has different characteristics from those discussed above.

This web of quadric cones contains ∞^3 nets of quadric cones. Each net has a jacobian cubic cone J_i whose vertex is at Q and whose plane section is an elliptic cubic. The ∞^3 J_i form a web. Two J_i intersect in nine generators of which three are the axes

of the three pairs of planes of the pencil of quadric cones common to the two nets. The other six generators are common to all the ∞^3 J_4 and are thus basis lines of the web J_4.

The jacobian of a net of jacobian cones J_4 is a cone of order 6, consisting of a quadric cone of the web and four planes through Q and through the six basis lines of the jacobian web. The four planes and six lines through Q form a complete four-plane.

The (1, 1) correspondence between the web of cones and the planes of (y) does not determine an involution between the points of the two spaces. There is, however, a surface L in (y) whose image is the bundle of lines on Q. The bundle of lines on Q may be considered the jacobian of the web of cones since any line through Q may be the axis of a composite cone of the web. To a plane section of L corresponds a quadric cone of the web.

The surface L in (y) is a Steiner's quartic. It has the characteristics: $N=4$, $n'=3$, $a=a'=6$, $b=3$, $c=0$, $\delta'=4$, $\kappa'=6$, $\kappa=9$, $\delta=0$, $t=1$, $C'=4$, j (pinch points of b) $=6$.

The nodal curve b consists of three nodal lines through one point, which is the triple point of b and of L. There are two pinch points on each nodal line.

The images of the six pinch points are the six basis lines of the web of jacobian cones. There are six pencils of quadric cones of the web in each of which four generators of every quadric cone of the pencil are consecutive with one basis line.

The images of the four contact conics of the four conic tropes of L are, respectively, the four planes through the six basis lines and Q. These four planes are the loci of the limiting positions of axes of composite cones whose planes coincide.

The image of b is a cubic cone through the six basic lines of the jacobian web and with three nodal generators corresponding to the triple point on b. This cubic cone consists of three planes through Q and is the diagonal trihedral of the basic complete four-plane. These three planes constitute the locus of pairs of lines through Q such that the quadric cones of a pencil have contact along the lines of a pair. In the web, there is one and only one pencil of quadric cones in which the cones have contact along three lines. These lines are the edges of the diagonal trihedral.

The web of quadric cones with one basic generator l gives

rise to a web of jacobian cubic cones with l as a common nodal generator and two other basis lines. The jacobian of a net of jacobian cones consists of a quadric cone of the web and two planes, each counted twice, both through l and each through one of the two basis lines of the jacobian web. In (y), L is a ruled cubic with the characteristics: $N=3$, $n'=3$, $a=a'=4$, $b=1$, $c=0$, $\kappa'=3$, $\delta'=0$, $\kappa=3$, $\delta=0$, $t=0$, $j=2$, $C'=2$. The image of b is a plane through the two basis lines, but not through l. The images of the two pinch points of b are the two basis lines. The images of the two conic tropes are the two planes determined by l and the two basis lines paired with l.

The web of quadric cones with two basis generators l_1 and l_2 gives a web of jacobian cubic cones, each consisting of the plane $l_1 l_2$ and a quadric cone of the web. The jacobian of a net of these composite cubic cones consists of a quadric cone of the web and the plane $l_1\, l_2$ counted four times. In (y), the surface L is a quadric. The web has contacts only in the plane $l_1\, l_2$.

The web of quadric cones can have no more than two basis lines or two basis points. One or two basis points determine one or two basis generators, respectively, and thus the discussion of basis points reduces to that of the two preceding cases.

WELLS COLLEGE

ERRATUM

On page 692 of the October issue of this Bulletin, in lines 5 and 19, change k to K.